B3

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 July 2003 (17.07.2003)

PCT

(10) International Publication Number
WO 03/057927 A2

(51) International Patent Classification[7]: C12Q 1/70

(21) International Application Number: PCT/GB03/00030

(22) International Filing Date: 7 January 2003 (07.01.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0200258.2 7 January 2002 (07.01.2002) GB

(71) Applicant (*for all designated States except US*): **NORCHIP A/S** [NO/NO]; Industriveien 8, N-3490 Klokkarstua (NO).

(71) Applicant (*for MG only*): **ALLARD, Susan, Joyce** [GB/GB]; Verulam Gardens, 70 Gray's Inn Road, London WC1X 8BT (GB).

(72) Inventor; and
(75) Inventor/Applicant (*for US only*): **KARLSEN, Frank** [NO/NO]; Hovkastet 2, N-3490 Klokkarstua (NO).

(74) Agents: **BALDOCK, Sharon, Claire** et al.; Boult Wade Tennant, Verulam Gardens, 70 Gray's Inn Road, London WC1X 8BT (GB).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv)) for US only*

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/057927 A2

(54) Title: DETECTION OF HUMAN PAPILLOMAVIRUS E6 MRNA

(57) Abstract: An oligonucleotide molecule for use in the detection of mRNA transcribed from the E6 gene of a human papillomavirus, the oligonucleotide comprising any one of sequence numbers 1-133.

DETECTION OF HUMAN PAPILLOMAVIRUS E6 mRNA

The present invention is concerned with oligonucleotide primers and probes for use in detecting the presence of mRNA transcripts from the E6 gene of human papillomavirus in clinical samples.

In the last few years, there has been an improvement in the methods used to detect HPV, with methods based on amplification of nucleic acids using the polymerase chain reaction (PCR) becoming increasingly widespread. It is now possible to detect small amounts of HPV DNA (<100 pg), quantify the amount of viral DNA in clinical samples, identify a broad spectrum of genital HPV types, test for selected HPV types and localise the viral genome transcripts and proteins to the individual cells. Since HPV detection is often carried out in the presence of vast quantities of host nucleic acids and cells not infected with the virus, the ability of the primers to be virus specific is critical for a sensitive and specific amplification.

The present inventors have selected new primer and probe sequences, specific for the E6 region, which may be used in the detection of E6 transcripts by the NASBA technique, particularly sensitive, real-time NASBA, or by RT-PCR. The inventors' approach is based upon the development of primers specific for regions of E6 which are conserved across high-risk, cancer-associated HPV types.

Therefore, in accordance with a first aspect the invention provides target-specific primers and oligonucleotide probes for use in the detection of human papillomavirus (HPV) E6 mRNA, particularly for use in detection of HPV E6 mRNA by RT-PCR or NASBA.

In particular, the invention provides primer and probe oligonucleotides comprising the HPV-specific sequences represented as sequence numbers (SEQ NO) 1 to 133 in Table 1. For each individual sequence an indication is given in the column "primer/probe type" of the general types of primers or probes into which the HPV-specific sequence may be incorporated for the purposes of HPV detection. The HPV type and position in the HPV genome is also indicated.

Table 1-Summary of primer sequences

PRIMER/PROBE TYPE	SEQUENCE	SEQ NO	HPV	nt
NASBA P2/PCR	CCACAGGAGCGACCCAGAAAGTTA	1	16	116
NASBA P1/PCR	ACGGTTTGTTGTATTGCTGTTC	2	16	368
NASBA P2/PCR	CCACAGGAGCGACCCAGAAA	3	16	116
NASBA P1/PCR	GGTTTGTTGTATTGCTGTTC	4	16	368
NASBA P1/PCR	TCACGTCGCAGTAACTGT	126	16	208
NASBA P1/PCR	TTGCTTGCAGTACACACA	127	16	191
NASBA P1/PCR	TGCAGTACACACATTCTA	128	16	186
NASBA P1/PCR	GCAGTACACACATTCTAA	129	16	185
NASBA P2/PCR PROBE	ACAGTTATGCACAGAGCT	130	16	142
NASBA P2/PCR PROBE	ATATTAGAATGTGTGTAC	131	16	182
NASBA P2/PCR PROBE	TTAGAATGTGTGTACTGC	132	16	185
NASBA P2/PCR PROBE	AATGTGTGTACTGCAAG	133	16	188
PROBE	CTTTGCTTTTCGGGATTTATGC	5	16	235
PROBE	TATGACTTTGCTTTTCGGGA	6	16	230
NASBA P2/PCR	CAGAGGAGGAGGATGAAATAGTA	7	16	656
NASBA P1/PCR	GCACAACCGAAGCGTAGAGTCACAC	8	16	741
PROBE	TGGACAAGCAGAACCGGACAGAGC	9	16	687
NASBA P2/PCR	CAGAGGAGGAGGATGAAATAGA	10	16	656
NASBA P1/PCR	GCACAACCGAAGCGTAGAGTCA	11	16	741
PROBE	AGCAGAACCGGACAGAGCCCATTA	12	16	693
NASBA P2/PCR	ACGATGAAATAGATGGAGTT	13	18	702
NASBA P1/PCR	CACGGACACACAAAGGACAG	14	18	869
PROBE	AGCCGAACCACAACGTCACA	15	18	748
NASBA P2/PCR	GAAAACGATGAAATAGATGGAG	16	18	698
NASBA P1/PCR	ACACCACGGACACACAAAGGACAG	17	18	869
PROBE	GAACCACAACGTCACACAATG	18	18	752
NASBA P2/PCR	TTCCGGTTGACCTTCTATGT	19	18	651
NASBA P1/PCR	GGTCGTCTGCTGAGCTTTCT	20	18	817
NASBA P2/PCR	GCAAGACATAGAAATAACCTG	21	18	179

NASBA P1/PCR	ACCCAGTGTTAGTTAGTT	22	18	379
PROBE	TGCAAGACAGTATTGGAACT	23	18	207
NASBA P2/PCR	GGAAATACCCTACGATGAAC	24	31	164
NASBA P1/PCR	GGACACAACGGTCTTTGACA	25	31	423
PROBE	ATAGGGACGACACACCACACGGAG	26	31	268
NASBA P2/PCR	GGAAATACCCTACGATGAACTA	27	31	164
NASBA P1/PCR	CTGGACACAACGGTCTTTGACA	28	31	423
PROBE	TAGGGACGACACACCACACGGA	29	31	269
NASBA P2/PCR	ACTGACCTCCACTGTTATGA	30	31	617
NASBA P1/PCR	TATCTACTTGTGTGCTCTGT	31	31	766
PROBE	GACAAGCAGAACCGGACACATC	32	31	687
NASBA P2/PCR	TGACCTCCACTGTTATGAGCAATT	33	31	619
NASBA P1/PCR	TGCGAATATCTACTTGTGTGCTCTGT	34	31	766
PROBE	GGACAAGCAGAACCGGACACATCCAA	35	31	686
NASBA P2/PCR	ACTGACCTCCACTGTTAT	36	31	617
NASBA P1/PCR	CACGATTCCAAATGAGCCCAT	37	31	809
NASBA P2/PCR	TATCCTGAACCAACTGACCTAT	38	33	618
NASBA P1/PCR	TTGACACATAAACGAACTG	39	33	763
PROBE	CAGATGGACAAGCACAACC	40	33	694
NASBA P2/PCR	TCCTGAACCAACTGACCTAT	41	33	620
NASBA P1/PCR	CCCATAAGTAGTTGCTGTAT	42	33	807
PROBE	GGACAAGCACAACCAGCCACAGC	43	33	699
NASBA P2/PCR	GACCTTTGTGTCCTCAAGAA	44	33	431
NASBA P1/PCR	AGGTCAGTTGGTTCAGGATA	45	33	618
PROBE	AGAAACTGCACTGTGACGTGT	46	33	543
NASBA P2/PCR	ATTACAGCGGAGTGAGGTAT	47	35	217
NASBA P1/PCR	GTCTTTGCTTTTCAACTGGA	48	35	442
NASBA P2/PCR	TCAGAGGAGGAGGAAGATACTA	49	35	655
NASBA P1/PCR	GATTATGCTCTCTGTGAACA	50	35	844
NASBA P2/PCR	CCCGAGGCAACTGACCTATA	51	35	610
NASBA P1/PCR	GTCAATGTGTGTGCTCTGTA	52	35	770
PROBE	ATAGAGAAGGCCAGCCATAT	53	35	270
PROBE	GACAAGCAAAACCAGACACCTCCAA	54	35	692
PROBE	GACAAGCAAAACCAGACACC	55	35	692
NASBA P2/PCR	TTGTGTGAGGTGCTGGAAGAAT	56	52	144
NASBA P1/PCR	CCCTCTCTTCTAATGTTT	57	52	358
PROBE	GTGCCTACGCTTTTTATCTA	58	52	296
NASBA P2/PCR	GTGCCTACGCTTTTTATCTA	59	52	296
NASBA P1/PCR	GGGGTCTCCAACACTCTGAACA	60	52	507
PROBE	TGCAAACAAGCGATTTCA	61	52	461
NASBA P2/PCR	TCAGGCGTTGGAGACATC	62	58	157
NASBA P1/PCR	AGCAATCGTAAGCACACT	63	58	301
NASBA P2/PCR	TCTGTGCATGAAATCGAA	64	58	173
NASBA P1/PCR	AGCACACTTTACATACTG	65	58	291
PROBE	TGAAATGCGTTGAATGCA	66	58	192
PROBE	TTGCAGCGATCTGAGGTATATG	67	58	218
NASBA P2/PCR	TACACTGCTGGACAACAT	68	B	514
NASBA P1/PCR	TCATCTTCTGAGCTGTCT	69	B	619
NASBA P2/PCR	TACACTGCTGGACAACATGCA	70	B	514
NASBA P1/PCR	GTCACATCCACAGCAACAGGTCA	71	B	693
PROBE	GTAGGGTTACATTGCTATGA	72	B	590
PROBE	GTAGGGTTACATTGCTATGAGC	73	B	590
NASBA P2/PCR	TGACCTGTTGCTGTGGATGTGA	74	B	693

NASBA P1/PCR	TACCTGAATCGTCCGCCAT	75	B	832
PROBE	ATWGTGTGTCCCATCTGC	76	B	794
NASBA P2/PCR	CATGCCATAAATGTATAGA	77	C	295
NASBA P1/PCR	CACCGCAGGCACCTTATTAA	78	C	408
PROBE	AGAATTAGAGAATTAAGA	79	C	324
NASBA P2/PCR	GCAGACGACCACTACAGCAAA	80	39	210
NASBA P1/PCR	ACACCGAGTCCGAGTAATA	81	39	344
PROBE	ATAGGGACGGGGAACCACT	82	39	273
NASBA P2/PCR	TATTACTCGGACTCGGTGT	83	39	344
NASBA P1/PCR	CTTGGGTTTCTCTTCGTGTTA	84	39	558
PROBE	GGACCACAAAACGGGAGGAC	85	39	531
NASBA P2/PCR	GAAATAGATGAACCCGACCA	86	39	703
NASBA P1/PCR	GCACACCACGGACACACAAA	87	39	886
PROBE	TAGCCAGACGGGATGAACCACAGC	88	39	749
NASBA P2/PCR	AACCATTGAACCCAGCAGAAA	89	45	430
NASBA P1/PCR	TCTTTCTTGCCGTGCCTGGTCA	90	45	527
NASBA P2/PCR	GAAACCATTGAACCCAGCAGAAAA	91	45	428
NASBA P1/PCR	TTGCTATACTTGTGTTTCCCTACG	92	45	558
PROBE	GTACCGAGGGCAGTGTAATA	93	45	500
PROBE	GGACAAACGAAGATTTCACA	94	45	467
NASBA P2/PCR	GTTGACCTGTTGTGTTACCAGCAAT	95	45	656
NASBA P1/PCR	CACCACGGACACACAAAGGACAAG	96	45	868
NASBA P2/PCR	CTGTTGACCTGTTGTGTTACGA	97	45	654
NASBA P1/PCR	CCACGGACACACAAAGGACAAG	98	45	868
NASBA P2/PCR	GTTGACCTGTTGTGTTACGA	99	45	656
NASBA P1/PCR	ACGGACACACAAAGGACAAG	100	45	868
PROBE	GAGTCAGAGGAGGAAAACGATG	101	45	686
PROBE	AGGAAAACGATGAAGCAGATGGAGT	102	45	696
PROBE	ACAACTACCAGCCCGACGAGCCGAA	103	45	730
NASBA P2/PCR	GGAGGAGGATGAAGTAGATA	104	51	658
NASBA P1/PCR	GCCCATTAACATCTGCTGTA	105	51	807
NASBA P2/PCR	AGAGGAGGAGGATGAAGTAGATA	106	51	655
NASBA P1/PCR	ACGGGCAAACCAGGCTTAGT	107	51	829
PROBE	GCAGGTGTTCAAGTGTAGTA	108	51	747
PROBE	TGGCAGTGGAAAGCAGTGGAGACA	109	51	771
NASBA P2/PCR	TTGGGGTGCTGGAGACAAACATCT	110	56	519
NASBA P1/PCR	TTCATCCTCATCCTCATCCTCTGA	111	56	665
NASBA P2/PCR	TGGGGTGCTGGAGACAAACATC	112	56	520
NASBA P1/PCR	CATCCTCATCCTCATCCTCTGA	113	56	665
NASBA P2/PCR	TTGGGGTGCTGGAGACAAACAT	114	56	519
NASBA P1/PCR	CCACAAACTTACACTCACAACA	115	56	764
PROBE	AAAGTACCAACGCTGCAAGACGT	116	56	581
PROBE	AGAACTAACACCTCAAACAGAAAT	117	56	610
PROBE	AGTACCAACGCTGCAAGACGTT	118	56	583
PROBE	TTGGACAGCTCAGAGGATGAGG	119	56	656
NASBA P2/PCR	GATTTTCCTTATGCAGTGTG	120	56	279
NASBA P1/PCR	GACATCTGTAGCACCTTATT	121	56	410
PROBE	GACTATTCAGTGTATGGAGC	122	56	348
PROBE	CAACTGAYCTMYACTGTTATGA	123	A	
PROBE	GAAMCAACTGACCTAYWCTGCTAT	124	A	
PROBE	AAGACATTATTCAGACTC	125	A	

Oligonucleotides for use as NASBA P1 primers have the general structure "X_1-SEQ", wherein "X_1" represents a nucleotide sequence comprising a promoter and "SEQ" represents the HPV-specific sequence, as given in Table 1. The inclusion of a promoter sequence is essential in NASBA P1 primers but is not necessary in PCR primers, as discussed below. In a preferred embodiment, X_1 may be a sequence comprising a bacteriophage promoter, preferably the T7 promoter. In the most preferred embodiment, X_1 represents the sequence AATTCTAATACGACTCACTATAGGGAGAAGG.

The oligonucleotide molecules of the invention are selected to be specific for mRNA transcribed from the HPV E6 gene. Active expression of the E7 and E6 genes of HPV is associated with cervical cytological abnormalities which often progress to more serious disease. A number of studies relate the expression of the E6 and E7 genes to oncogenesis. Co-operation between E6 and E7 increases significantly the frequency of immortalization. Evidence has been presented that the E6 and E7 open reading frames are involved in the transforming activity of the virus (Tanaka et al., J. Virol. 63: 1465-1469, 1989). These transformation effects of E6 and E7 may at least in part be explained by their interaction with the cellular tumour suppressor gene products p53 and pRb 105, respectively (Boyer et al., Cancer Research. 56: 4620-4624, 1996; Lechner et al. EMBO J. 11: 3045-3051, 1992).

HPV16 mRNA isolated from transfected cells and a variety of tumour cell lines and lesions containing both extrachromosomal and integrated HPV16 genomes has

been analysed in multiple laboratories (see Doorbar JA et al., Virology 178:254?262, Rohlfs et al., Virology 183:331?342; Sherman et al., Int. J. Cancer 50:356?364). These studies have shown that several different alternatively spliced transcripts may be produced from the E6 and E7 region. In summary, there are four major transcripts: one with the whole E6/E7 gene area (E6), one with a loss of a coding sequence between basepairs 226 and 409 (E6*I), one with a loss of a coding sequence in a larger part of E6 between 226 and 526 (E6*II) and one with the loss of the E7 transcript (E6*III). However there are clearly consensus sequences in the area up to 226 basepairs in the E6 region. The inventors therefore selected the areas between 97 and 226 and between 526 and 880 as areas to target for diagnostic purposes.

The oligonucleotides provided by the invention may be grouped according to specificity for different specific HPV types or groups of HPV types. Sequence numbers 1-12 and 126-133 are specific for HPV type 16, sequence numbers 13-23 are specific for HPV type 18, sequence numbers 24-37 are specific for HPV type 31, sequence numbers 38-46 are specific for HPV type 33. HPV types 16, 18 , 31 and 33 are the major cancer-associated HPV types. Sequence numbers 47-55 are specific for HPV type 35, sequence numbers 56-61 are specific for HPV type 52, sequence numbers 62-67 are specific for HPV type 58, sequence numbers 80-88 are specific for HPV type 39, sequence numbers 89-103 are specific for HPV type 45, sequence numbers 104-109 are specific for HPV type 51, sequence numbers 110-122 are specific for HPV type 56. Sequence numbers 68-76 are consensus sequences for group B HPV types (in

particular HPV types 6 and 11). Sequence numbers 77-79 and 125 are consensus sequences for group C HPV types (including HPV types 18, 39 and 45). Sequence numbers 123 and 124 are consensus probe sequences for group A HPV types. Sequence 123 is a consensus for HPV types 16, 31 and 35; sequence 124 is a consensus for HPV types 33, 52 and 58).

The oligonucleotide molecules of the invention are preferably single stranded DNA molecules. Non-natural synthetic polynucleotides which retain the ability to base-pair with a complementary nucleic acid molecule and are also within the scope of the invention, including synthetic oligonucleotides which incorporate modified bases and synthetic oligonucleotides wherein the links between individual nucleosides include bonds other than phosphodiester bonds. The oligonucleotide molecules of the invention may be produced according to techniques well known in the art, such as by chemical synthesis using standard apparatus and protocols for oligonucleotide synthesis.

The oligonucleotide molecules provided by the invention will typically be isolated single-stranded polynucleotides of no more than 100 bases in length, more typically less than 55 bases in length. For the avoidance of doubt it is hereby stated that the language "oligonucleotide comprising sequence number n" excludes the naturally occurring full-length HPV genomes.

The invention provides several general types of oligonucleotide primers and probes incorporating the HPV-specific sequences listed in Table 1. Typically,

such oligonucleotides may comprise additional, non-HPV sequences, for example sequences which are required for an amplification reaction or which facilitate detection of the products of the amplification reaction. The HPV-specific part of the oligonucleotide may consist of one of the sequences listed in Table 1 in the absence of any other contiguous HPV sequences. However, it will be appreciated that minor variations may be made to the HPV-specific sequences, for example the addition, deletion or substitution of bases, without affecting the ability of the oligonucleotide to bind to its target sequence and function as a primer or probe to a material extent.

The first type of oligonucleotides are primer 1 oligonucleotides (also referred to herein as NASBA P1 primers), which are oligonucleotides of generally approximately 50 bases in length, containing an average of about 20 bases at the 3' end that are complementary to a region of the target mRNA. Oligonucleotides suitable for use as NASBA P1 primers are denoted "NASBA P1/PCR" in Table 1. The 5' ends of the P1 primer oligonucleotides (represented herein in general terms as X_1) comprise a promoter sequence that is recognized by a specific RNA polymerase. Bacteriophage promoters, for example the T7, T3 and SP6 promoters, are preferred for use in the oligonucleotides of the invention, since they provide advantages of high level transcription which is dependent only on binding of the appropriate RNA polymerase. In a preferred embodiment, the 5' terminal sequence of the P1 primer oligonucleotides may comprise the sequence AATTCTAATACGACTCACTATAGGG or

the sequence AATTCTAATACGACTCACTATAGGGAGAAGG. These sequences contains a T7 promoter, including the transcription initiation site for T7 RNA polymerase. The HPV-specific sequences denoted in Table 1 as "NASBA P1/PCR" are suitable for use in both NASBA P1 primers and standard PCR primers. When these sequences are used as the basis of NASBA P1 primers they have the general structure X_1-SEQ, wherein X_1 represents a sequence comprising a promoter and SEQ represents the HPV-specific sequence. The promoter sequence X_1 is essential. However, when the same sequences are used as the basis of standard PCR primers it is not necessary to include X_1. The phrase "sequence number" as used in the claims is to be interpreted accordingly.

For the avoidance of doubt, the phrase "a NASBA P1 primer comprising sequence number 1" is to be interpreted as requiring the presence of an X_1 sequence 5' to the HPV-specific sequence listed as sequence number 1, whereas the phrase "a PCR primer comprising sequence number 1" refers to any suitable PCR primer comprising the HPV-specific sequence, X_1 not being an essential feature of a PCR primer. The phrase "an oligonucleotide primer including sequence number n" is taken to encompass NASBA P1, NASBA P2 and PCR primers.

A second type of oligonucleotide provided by the invention are NASBA primer 2 oligonucleotides (also referred to herein as NASBA P2 primers) which generally comprise a sequence of approximately 20 bases substantially identical to a region of the target mRNA. The oligonucleotide sequences denoted in

Table 1 as "NASBA P2/PCR" are suitable for use in both NASBA P1 primers and standard PCR primers.

Oligonucleotides intended for use as NASBA P2 primers may, in a particular but non-limiting embodiment, further comprise a sequence of nucleotides at the 5' end which is unrelated to the target mRNA but which is capable of hybridising to a generic detection probe. The detection probe will preferably be labelled, for example with a fluorescent, luminescent or enzymatic label. In one embodiment the detection probe is labelled with a label that permits detection using ECL™ technology, although it will be appreciated that the invention is in no way limited to this particular method of detection. In a preferred embodiment the 5' end of the primer 2 oligonucleotides may comprise the sequence GATGCAAGGTCGCATATGAG. This sequence is capable of hybridising to a generic ECL™ probe commercially available from Organon Teknika having the following structure:

$Ru(bpy)_3^{2+}$-GAT GCA AGG TCG CAT ATG AG-3'

In a different embodiment the primer 2 oligonucleotide may incorporate "molecular beacons" technology, which is known in the art and described, for example, in WO 95/13399 by Tyagi and Kramer, Nature Biotechnology. 14: 303-308, 1996, to allow for real-time monitoring of the NASBA reaction.

A third type of oligonucleotide molecules provided by the invention are target-specific probe oligonucleotides (denoted "probe" in Table 1). The probe oligonucleotides generally comprise a sequence of approximately 20-25 bases substantially identical

to a region of the target mRNA, or the complement thereof. The probe oligonucleotides may be used as target-specific hybridisation probes for detection of the products of a NASBA or PCR reaction. In this connection the probe oligonucleotides may be coupled to a solid support, such as paramagnetic beads, to form a capture probe (see below). In a preferred embodiment the 5' end of the probe oligonucleotide may be labelled with biotin. The addition of a biotin label facilitates attachment of the probe to a solid support via a biotin/streptavidin or biotin/avidin linkage.

A fourth type of oligonucleotide molecules provided by the invention are target-specific probes incorporating "molecular beacons" technology which is known in the art and described, for example, by Tyagi and Kramer, Nature Biotechnology. 14: 303-308, 1996 and in WO 95/13399.

The term "molecular beacons probes" as used herein is taken to mean molecules having the structure:

X_2-arm_1-target-arm_2-X_3

wherein "target" represents a target-specific sequence of nucleotides, "X_2" and "X_3" represent a fluorescent moiety and a quencher moiety capable of substantially or completely quenching the fluorescence from the fluorescent moiety when the two are held together in close proximity and "arm_1" and "arm_2" represent complementary sequences capable of forming a stem duplex.

The invention provides molecular beacons probes incorporating a target-specific sequence comprising one of sequence numbers 6, 18, 35, 43, 123, 124 or 125.

Suitable pairs of arm_1 and arm_2 sequences for use with these HPV-specific sequences include, but not exclusively, the following:

For use with sequence number 6:
CGCATG---------CATGCG
CCAGCT---------AGCTGG
CACGC-----------GCGTG
CGATCG---------CGATCG

For use with sequence number 18:
CGCATG---------CATGCG
CCGTCG---------CGACGG
CGGACC---------GGTCCG
CGATCG---------CGATCG

For use with sequence number 35:
CCGAAGG-------CCTTCGG
CCGTCG---------CGACGG
CACGTCG-------CGACGTG
CGCAGC---------GCTGCG
CGATCG---------CGATCG

For use with sequence number 43:
CCAAGC---------GCTTGG
CCAAGCG-------CGCTTGG
CCCAGC---------GCTGGG
CCAAAGC-------GCTTTGG
CCTGC-----------GCAGG

CGATCG----------CGATCG

For use with sequence number 123:

CGCATG----------CATGCG

CCGTCG----------CGACGG

CCACCC-----------GGGTGG

CGATCG-----------CGATCG

For use with sequence number 124:

CCAAGC----------GCTTGG

CCAAGCC---------GGCTTGG

CCAAGCG---------GCGTTGG

CCAGCG-----------CGCTGG

CGATCG-----------CGATCG

For use with sequence number 125:

CCAAGC-----------GCTTGG

CGCATG-----------CATGCG

CCCAGC-----------GCTGGG

CGATCG-----------CGATCG

The use of probe molecules incorporating molecular beacons technology allows for real-time monitoring of amplification reactions, such as NASBA or RT-PCR reactions. The use of molecular beacons technology allows for real-time monitoring of the NASBA reaction (see Leone et al., Nucleic Acids Research., 1998, vol: 26, pp 2150-2155). The molecular beacons probes generally include complementary sequences flanking the HPV-specific sequence, represented herein by the notation arm_1 and arm_2, which are capable of hybridising to each other form a stem duplex structure. The precise sequences of arm_1 and arm_2 are not material to the invention,

except for the requirement that these sequences must be capable of forming a stem duplex when the probe is not bound to a target HPV sequence.

Molecular beacons probes also include a fluorescent moiety and a quencher moiety, the fluorescent and the quencher moieties being represented herein by the notation X_2 and X_3. As will be appreciated be the skilled reader, the fluorescer and quencher moieties are selected such that the quencher moiety is capable of substantially or completely quenching the fluorescence from the fluorescent moiety when the two moieties are in close proximity, e.g. when the probe is in the hairpin "closed" conformation in the absence of the target sequence. Upon binding to the target sequence, the fluorescent and quencher moieties are held apart such that the fluorescence of the fluorescent moiety is no longer quenched.

Many examples of suitable pairs of quencher/fluorescer moieties which may be used in accordance with the invention are known in the art (see WO 95/13399, Tyagi and Kramer, ibid). A broad range of fluorophores in many different colours made be used, including for example 5-(2'-aminoethyl)aminonaphthalene-1-sulphonic acid (EDANS), fluorescein, FAM and Texas Red (see Tyagi, Bratu and Kramer, 1998, Nature Biotechnology, 16, 49-53. The use of probes labelled with different coloured fluorophores enables "multiplex" detection of two or more different probes in a single reaction vessel. A preferred quencher is 4-(4'-dimethylaminophenylazo)benzoic acid (DABCYL), a

non-fluorescent chromophore, which serves as a 'universal' quencher for a wide range of fluorophores. The fluorescer and quencher moieties may be covalently attached to the probe in either orientation, either with the fluorescer at or near the 5' end and the quencher at or near the 3' end or vice versa. Protocols for the synthesis of molecular beacon probes are known in the art. A detailed protocol for synthesis is provided in a paper entitled "Molecular Beacons: Hybridization Probes for Detection of Nucleic Acids in Homogenous Solutions" by Sanjay Tyagi et al., Department of Molecular Genetics, Public Health Research Institute, 455 First Avenue, New York, NY 10016, USA, which is available online via the PHRI website (at www.phri.nyu.edu or www.molecular-beacons.org).

Suitable combinations of the NASBA P1 and NASBA P2 primer oligonucleotide molecules provided by the invention may be used to drive a NASBA amplification reaction. In order to drive a NASBA amplification reaction the primer 1 and primer 2 oligonucleotides must be capable of priming synthesis of a double-stranded DNA from a target region of mRNA. For this to occur the primer 1 and primer 2 oligonucleotides must comprise target-specific sequences which are complementary to regions of the sense and the antisense strand of the target mRNA, respectively.

In the first phase of the NASBA amplification cycle, the so-called "non-cyclic" phase, the primer 1 oligonucleotide anneals to a complementary sequence in the target mRNA and its 3' end is extended by the

action of an RNA-dependent DNA polymerase (e.g. reverse transcriptase) to form a first-strand cDNA synthesis. The RNA strand of the resulting RNA:DNA hybrid is then digested, e.g. by the action of RNaseH, to leave a single stranded DNA. The primer 2 oligonucleotide anneals to a complementary sequence towards the 3' end of this single stranded DNA and its 3' end is extended (by the action of reverse transcriptase), forming a double stranded DNA. RNA polymerase is then able to transcribe multiple RNA copies from the now transcriptionally active promoter sequence within the double-stranded DNA. This RNA transcript, which is antisense to the original target mRNA, can act as a template for a further round of NASBA reactions, with primer 2 annealing to the RNA and priming synthesis of the first cDNA strand and primer 1 priming synthesis of the second cDNA strand. The general principles of the NASBA reaction are well known in the art (see Compton, J. Nature. 350: 91-92).

The target-specific probe oligonucleotides described herein may also be attached to a solid support, such as magnetic microbeads, and used as "capture probes" to immobilise the product of the NASBA amplification reaction (a single stranded RNA). The target-specific "molecular beacons" probes described herein may be used for real-time monitoring of the NASBA reaction.

In a particular embodiment the invention provides the oligonucleotide listed in Table 2, these being NASBA P1 primers and NASBA P2 primers containing the sequences listed in Table 1. The NASBA P1 primers further include a T7 promoter sequence, the NASBA P2

primers include a sequence for binding of a generic detection probe (see below) and associated probe molecules for use in the detection of HPV mRNA by NASBA. The oligonucleotides listed in Table 2 are merely illustrative and it is not intended that the scope of the invention should be limited to these specific molecules.

The NASBA P2 primers (p2)in Table 2 include the sequence GATGCAAGGTCGCATATGAG at the 5' end; the NASBA P1 primers (p1) in Table 2 include the sequence AATTCTAATACGACTCACTATAGGGAGAAGG at the 5' end. Oligonucleotides suitable for use as probes are identified by "po". The P2 primers generally contain HPV sequences from the postive strand, whereas the p1 primers generally contain HPV sequences from the negative strand. nt-refers to nucleotide position in the relevant HPV genomic sequence.

Table 2-NASBA primer and probe sequences

Primer name	**Sequence**	HPV Type	nt
HAe6701p2	GATGCAAGGTCGCATATGAGCCACAGGAGCGACCC AGAAAGTTA	16	116
HAe6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGACGG TTTGTTGTATTGCTGTTC	16	368
HAe6702p2	GATGCAAGGTCGCATATGAGCCACAGGAGCGACCC AGAAA	16	116
HAe6702p1	AATTCTAATACGACTCACTATAGGGAGAAGGGGTT TGTTGTATTGCTGTTC	16	368
HAe6702Ap1	AATTCTAATACGACTCACTATAGGGAGAAGGTCA CGTCGCAGTAACTGT	16	208
HAe6702Bp1	AATTCTAATACGACTCACTATAGGGAGAAGGTTG CTTGCAGTACACACA	16	191
HAe6702Cp1	AATTCTAATACGACTCACTATAGGGAGAAGGTGC AGTACACACATTCTA	16	186
HAe6702Dp1	AATTCTAATACGACTCACTATAGGGAGAAGGGCA GTACACACATTCTAA	16	185
H16e6702Ap2	GATGCAAGGTCGCATATGAGACAGTTATGCACAGA GCT	16	142

Primer name	Sequence	HPV Type	nt
H16e6702Bp2	GATGCAAGGTCGCATATGAGATATTAGAATGTGTGTAC	16	182
H16e6702Cp2	GATGCAAGGTCGCATATGAGTTAGAATGTGTGTACTGC	16	185
H16e6702Dp2	GATGCAAGGTCGCATATGAGGAATGTGTGTACTGCAAG	16	188
H16e6702Apo	ACAGTTATGCACAGAGCT	16	142
H16e6702Bpo	ATATTAGAATGTGTGTAC	16	182
H16e6702Cpo	TTAGAATGTGTGTACTGC	16	185
H16e6702Dpo	GAATGTGTGTACTGCAAG	16	188
HAe6701po	CTTTGCTTTTCGGGATTTATGC	16	235
HAe6702po	TATGACTTTGCTTTTCGGGA	16	230
HAe6702mb1	X_2-cgcatgTATGACTTTGCTTTTCGGGAcatgcg-X_3	16	230
HAe6702mb2	X_2-ccagctTATGACTTTGCTTTTCGGGAagctgg-X_3	16	230
HAe6702mb3	X_2-cacgcTATGACTTTGCTTTTCGGGAgcgtg-X_3	16	230
H16e6702mb4	X_2-cgatcgTATGACTTTGCTTTTCGGGAcgatcg-X_3	16	230
HAe6703p2	GATGCAAGGTCGCATATGAGCAGAGGAGGAGGATGAAATAGTA	16	656
HAe6703p1	AATTCTAATACGACTCACTATAGGGAGAAGGGCACAACCGAAGCGTAGAGTCACAC	16	741
HAe6703po	TGGACAAGCAGAACCGGACAGAGC	16	687
HAe6704p2	GATGCAAGGTCGCATATGAGCAGAGGAGGAGGATGAAATAGA	16	656
HAe6704p1	AATTCTAATACGACTCACTATAGGGAGAAGGGCACAACCGAAGCGTAGAGTCA	16	741
HAe6704po	AGCAGAACCGGACAGAGCCCATTA	16	693
H18e6701p2	GATGCAAGGTCGCATATGAGACGATGAAATAGATGGAGTT	18	702
H18e6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGCACGGACACACAAAGGACAG	18	869
H18e6701po	AGCCGAACCACAACGTCACA	18	748
H18e6702p2	GATGCAAGGTCGCATATGAGGAAAACGATGAAATAGATGGAG	18	698
H18e6702p1	AATTCTAATACGACTCACTATAGGGAGAAGGACACCACGGACACACAAAGGACAG	18	869
H18e6702po	GAACCACAACGTCACACAATG	18	752
H18e6702mb1	X_2-cgcatgGAACCACAACGTCACACAATGcatgcg-X_3	18	752
H18e6702mb2	X_2-ccgtcgGAACCACAACGTCACACAATGcgacgg-X_3	18	752
H18e6702mb3	X_2-cggaccGAACCACAACGTCACACAATGggtccg-X_3	18	752
H18e6702mb4	X_2-cgatcgGAACCACAACGTCACACAATGcgatcg	18	752

Primer name	Sequence	HPV Type	nt
	$-X_3$		
H18e6703p2	GATGCAAGGTCGCATATGAGTTCCGGTTGACCTTC TATGT	18	651
H18e6703p1	AATTCTAATACGACTCACTATAGGGAGAAGGGGTC GTCTGCTGAGCTTTCT	18	817
H18e6704p2	GATGCAAGGTCGCATATGAGGCAAGACATAGAAAT AACCTG	18	179
H18e6704p1	AATTCTAATACGACTCACTATAGGGAGAAGGACCC AGTGTTAGTTAGTT	18	379
H18e6704po	TGCAAGACAGTATTGGAACT	18	207
H31e6701p2	GATGCAAGGTCGCATATGAGGGAAATACCCTACGA TGAAC	31	164
H31e6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGGGAC ACAACGGTCTTTGACA	31	423
H31e6701po	ATAGGGACGACACACCACACGGAG	31	268
H31e6702p2	GATGCAAGGTCGCATATGAGGGAAATACCCTACGA TGAACTA	31	164
H31e6702p1	AATTCTAATACGACTCACTATAGGGAGAAGGCTGG ACACAACGGTCTTTGACA	31	423
H31e6702po	TAGGGACGACACACCACACGGA	31	269
H31e6703p2	GATGCAAGGTCGCATATGAGACTGACCTCCACTGT TATGA	31	617
H31e6703p1	AATTCTAATACGACTCACTATAGGGAGAAGGTATC TACTTGTGTGCTCTGT	31	766
H31e6703po	GACAAGCAGAACCGGACACATC	31	687
H31e6704p2	GATGCAAGGTCGCATATGAGTGACCTCCACTGTTA TGAGCAATT	31	619
H31e6704p1	AATTCTAATACGACTCACTATAGGGAGAAGGTGCG AATATCTACTTGTGTGCTCT GT	31	766
H31e6704po	GGACAAGCAGAACCGGACACATCCAA	31	686
H31e6704mb1	X_2-ccgaaggGGACAAGCAGAACCGGACACATCC AAccttcgg $-X_3$	31	686
H31e6704mb2	X_2-ccgtcgGGACAAGCAGAACCGGACACATCCA Acgacgg $-X_3$	31	686
H31e6704mb3	X_2- cacgtcgGGACAAGCAGAACCGGACACATCCAA cgacgtg $-X_3$	31	686
H31e6704mb4	X_2-cgcagcGGACAAGCAGAACCGGACACATCCAA gctgcg $-X_3$	31	686
H31e6704mb5	X_2-cgatcgGGACAAGCAGAACCGGACACATCCAA cgatcg $-X_3$	31	686
H31e6705p2	GATGCAAGGTCGCATATGAGACTGACCTCCACTGT TAT	31	617
H31e6705p1	AATTCTAATACGACTCACTATAGGGAGAAGGCACG ATTCCAAATGAGCCCAT	31	809
H33e6701p2	GATGCAAGGTCGCATATGAGTATCCTGAACCAACT GACCTAT	33	618

Primer name	Sequence	HPV Type	nt
H33e6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGTTGA CACATAAACGAACTG	33	763
H33e6701po	CAGATGGACAAGCACAACC	33	694
H33e6703p2	GATGCAAGGTCGCATATGAGTCCTGAACCAACTGA CCTAT	33	620
H33e6703p1	AATTCTAATACGACTCACTATAGGGAGAAGGCCCA TAAGTAGTTGCTGTAT	33	807
H33e6703po	GGACAAGCACAACCAGCCACAGC	33	699
H33e6703mb1	X_2-ccaagcGGACAAGCACAACCAGCCACAGCgct tgg -X_3	33	699
H33e6703mb2	X_2-ccaagcgGGACAAGCACAACCAGCCACAGC cgcttgg -X_3	33	699
H33e6703mb3	X_2-cccagcGGACAAGCACAACCAGCCACAGCgct ggg -X_3	33	699
H33e6703mb4	X_2-ccaaagcGGACAAGCACAACCAGCCACAGCg ctttgg -X_3	33	699
H33e6703mb5	X_2-cctgcGGACAAGCACAACCAGCCACAGCgcagg -X_3	33	699
H33e6703mb6	X_2-cgatcgGGACAAGCACAACCAGCCACAGCcga tcg -X_3	33	699
H33e6702p2	GATGCAAGGTCGCATATGAGGACCTTTGTGTCCTC AAGAA	33	431
H33e6702p1	AATTCTAATACGACTCACTATAGGGAGAAGGAGGT CAGTTGGTTCAGGATA	33	618
H33e6702po	AGAAACTGCACTGTGACGTGT	33	543
H35e6701p2	GATGCAAGGTCGCATATGAGATTACAGCGGAGTGA GGTAT	35	217
H35e6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGGTCT TTGCTTTTCAACTGGA	35	442
H35e5601po	ATAGAGAAGGCCAGCCATAT	35	270
H35e6702p2	GATGCAAGGTCGCATATGAGTCAGAGGAGGAGGAA GATACTA	35	655
H35e6702p1	AATTCTAATACGACTCACTATAGGGAGAAGGGATT ATGCTCTCTGTGAACA	35	844
H35e6703p2	GATGCAAGGTCGCATATGAGCCCGAGGCAACTGAC CTATA	35	610
H35e6703p1	AATTCTAATACGACTCACTATAGGGAGAAGGGTCA ATGTGTGTGCTCTGTA	35	770
H35e6702po	GACAAGCAAAACCAGACACCTCCAA	35	692
H35e6703po	GACAAGCAAAACCAGACACC	35	692
H52e6701p2	GATGCAAGGTCGCATATGAGTTGTGTGAGGTGCTG GAAGAAT	52	144
H52e6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGCCCT CTCTTCTAATGTTT	52	358
H52e6701po	GTGCCTACGCTTTTTATCTA	52	296
H52e6702p2	GATGCAAGGTCGCATATGAGGTGCCTACGCTTTTT	52	296

Primer name	Sequence	HPV Type	nt
	ATCTA		
H52e6702p1	AATTCTAATACGACTCACTATAGGGAGAAGGGGGG TCTCCAACACTCTGAACA	52	507
H52e6702po	TGCAAACAAGCGATTTCA	52	461
H58e6701p2	GATGCAAGGTCGCATATGAGTCAGGCGTTGGAGAC ATC	58	157
H58e6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGAGCA ATCGTAAGCACACT	58	301
H58e6702p2	GATGCAAGGTCGCATATGAGTCTGTGCATGAAATC GAA	58	173
H58e6702p1	AATTCTAATACGACTCACTATAGGGAGAAGGAGCA CACTTTACATACTG	58	291
H58e6701po	TGAAATGCGTTGAATGCA	58	192
H58e6702po	TTGCAGCGATCTGAGGTATATG	58	218
HBe6701p2	GATGCAAGGTCGCATATGAGTACACTGCTGGACAA CAT	B(11)	514
HBe6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGTCAT CTTCTGAGCTGTCT	B(11)	619
HBe6702p2	GATGCAAGGTCGCATATGAGTACACTGCTGGACAA CATGCA	B(11)	514
HBe6702p1	AATTCTAATACGACTCACTATAGGGAGAAGGGTCA CATCCACAGCAACAGGTCA	B(11)	693
HBe6701po	GTAGGGTTACATTGCTATGA	B(11)	590
HBe6702po	GTAGGGTTACATTGCTATGAGC	B(11)	590
HBe6703p2	GATGCAAGGTCGCATATGAGTGACCTGTTGCTGTG GATGTGA	B(11)	693
HBe6703p1	AATTCTAATACGACTCACTATAGGGAGAAGGTACC TGAATCGTCCGCCAT	B(11)	832
HBe6703po	ATWGTGTGTCCCATCTGC	B(11)	794
HCe6701p2	GATGCAAGGTCGCATATGAGCATGCCATAAATGTA TAGA	C(18 39 45)	295
HCe6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGCACC GCAGGCACCTTATTAA	C(18 39 45	408
HCe6701po	AGAATTAGAGAATTAAGA	C(18 39 45	324
H39e6701p2	GATGCAAGGTCGCATATGAGGCAGACGACCACTAC AGCAAA	39	210
H39e6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGACAC CGAGTCCGAGTAATA	39	344
H39e6701po	ATAGGGACGGGGAACCACT	39	273
H39e6702p2	GATGCAAGGTCGCATATGAGTATTACTCGGACTCG GTGT	39	344
H39e6702p1	AATTCTAATACGACTCACTATAGGGAGAAGGCTTG GGTTTCTCTTCGTGTTA	39	558
H39e6702po	GGACCACAAAACGGGAGGAC	39	531
H39e6703p2	GATGCAAGGTCGCATATGAGGAAATAGATGAACCC GACCA	39	703

Primer name	Sequence	HPV Type	nt
H39e6703p1	AATTCTAATACGACTCACTATAGGGAGAAGGGCAC ACCACGGACACACAAA	39	886
H39e6703po	TAGCCAGACGGGATGAACCACAGC	39	749
H45e6701p2	GATGCAAGGTCGCATATGAGAACCATTGAACCCAG CAGAAA	45	430
H45e6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGTCTT TCTTGCCGTGCCTGGTCA	45	527
H45e6702p2	GATGCAAGGTCGCATATGAGGAAACCATTGAACCC AGCAGAAAA	45	428
H45e6702p1	AATTCTAATACGACTCACTATAGGGAGAAGGTTGC TATACTTGTGTTTCCCTACG	45	558
H45e6701po	GTACCGAGGGCAGTGTAATA	45	500
H45e6702po	GGACAAACGAAGATTTCACA	45	467
H45e6703p2	GATGCAAGGTCGCATATGAGGTTGACCTGTTGTGT TACCAGCAAT	45	656
H45e6703p1	AATTCTAATACGACTCACTATAGGGAGAAGGCACC ACGGACACACAAAGGACAAG	45	868
H45e6704p2	GATGCAAGGTCGCATATGAGCTGTTGACCTGTTGT GTTACGA	45	654
H45e6704p1	AATTCTAATACGACTCACTATAGGGAGAAGGCCAC GGACACACAAAGGACAAG	45	868
H45e6705p2	GATGCAAGGTCGCATATGAGGTTGACCTGTTGTGT TACGA	45	656
H45e6705p1	AATTCTAATACGACTCACTATAGGGAGAAGGACGG ACACACAAAGGACAAG	45	868
H45e6703po	GAGTCAGAGGAGGAAAACGATG	45	686
H45e6704po	AGGAAAACGATGAAGCAGATGGAGT	45	696
H45e6705po	ACAACTACCAGCCCGACGAGCCGAA	45	730
H51e6701p2	GATGCAAGGTCGCATATGAGGGAGGAGGATGAAGT AGATA	51	658
H51e6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGGCCC ATTAACATCTGCTGTA	51	807
H51e6702p2	GATGCAAGGTCGCATATGAGAGAGGAGGAGGATGA AGTAGATA	51	655
H51e6702p1	AATTCTAATACGACTCACTATAGGGAGAAGGACGG GCAAACCAGGCTTAGT	51	829
H51e6701po	GCAGGTGTTCAAGTGTAGTA	51	747
H51e6702po	TGGCAGTGGAAAGCAGTGGAGACA	51	771
H56e6701p2	GATGCAAGGTCGCATATGAGTTGGGGTGCTGGAGA CAAACATCT	56	519
H56e6701p1	AATTCTAATACGACTCACTATAGGGAGAAGGTTCA TCCTCATCCTCATCCTCTGA	56	665
H56e6702p2	GATGCAAGGTCGCATATGAGTGGGGTGCTGGAGAC AAACATC	56	520
H56e6702p1	AATTCTAATACGACTCACTATAGGGAGAAGGCATC CTCATCCTCATCCTCTGA	56	665
H56e6703p2	GATGCAAGGTCGCATATGAGTTGGGGTGCTGGAGA	56	519

Primer name	Sequence	HPV Type	nt
	CAAACAT		
H56e6703p1	AATTCTAATACGACTCACTATAGGGAGAAGGCCACAAACTTACACTCACAACA	56	764
H56e6701po	AAAGTACCAACGCTGCAAGACGT	56	581
H56e6702po	AGAACTAACACCTCAAACAGAAAT	56	610
H56e6703po	AGTACCAACGCTGCAAGACGTT	56	583
H56e6703po1	TTGGACAGCTCAGAGGATGAGG	56	656
H56e6704p2	GATGCAAGGTCGCATATGAGGATTTTCCTTATGCAGTGTG	56	279
H56e6704p1	AATTCTAATACGACTCACTATAGGGAGAAGGGACATCTGTAGCACCTTATT	56	410
H56e6704po	GACTATTCAGTGTATGGAGC	56	348
HPVAPO1A	CAACTGAYCTMYACTGTTATGA	A (16 31 35)	
HPVApo1Amb1	X_2-cgcatgCAACTGAYCTMYACTGTTATGAcatgcg -X_3	A (16 31 35)	
HPVApo1Amb2	X_2-ccgtcgCAACTGAYCTMYACTGTTATGAcgacgg -X_3	A (16 31 35)	
HPVApo1Amb3	X_2-ccacccCAACTGAYCTMYACTGTTATGAgggtgg -X_3	A (16 31 35)	
HPVApo1Amb4	X_2-cgatcgCAACTGAYCTMYACTGTTATGAcgatcg -X_3	A (16 31 35)	
HPVAPO4A	GAAMCAACTGACCTAYWCTGCTAT	A (33 52 58)	
HPVAPO4Amb1	X_2-ccaagcGAAMCAACTGACCTAYWCTGCTATgcttgg -X_3	A (33 52 58)	
HPVAPO4Amb2	X_2-ccaagccGAAMCAACTGACCTAYWCTGCTATggcttgg -X_3	A (33 52 58)	
HPVAPO4Amb3	X_2-ccaagcgGAAMCAACTGACCTAYWCTGCTATcgcttgg -X_3	A (33 52 58)	
HPVAPO4Amb4	X_2-ccagcgGAAMCAACTGACCTAYWCTGCTATcgctgg -X_3	A (33 52 58)	
HPVAPO4Amb5	X_2-cgatcgGAAMCAACTGACCTAYWCTGCTATcgatcg -X_3	A (33 52 58)	
HPVCPO4	AAGACATTATTCAGACTC	C (18 45 39)	
HPVCPO4Amb1	X_2-ccaagcAAGACATTATTCAGACTCgcttgg -X_3	C (18 45 39)	
HPVCPO4Amb2	X_2-cgcatgAAGACATTATTCAGACTCcatgcg -X_3	C (18 45 39)	
HPVCPO4Amb3	X_2-cccagcAAGACATTATTCAGACTCgctggg -X_3	C (18 45 39)	
HPVCPO4Amb4	X_2-cgatcgAAGACATTATTCAGACTCcgatcg -X_3	C (18 45 39)	

The meaning of X_2- and -X_3 is defined above, in the discussion of "molecular beacons" probe molecules.

In a further embodiment the invention provides the oligonucleotides listed in Table 3, these being PCR primers for use in the detection of HPV mRNA by RT-PCR. These oligonucleotides are merely illustrative and it is not intended that the scope of the invention should be limited to these specific molecules:

Primer name	Sequence	HPV type	nt
HAe6701PCR2	CCACAGGAGCGACCCAGAAAGTTA	16	116
HAe6701PCR1	ACGGTTTGTTGTATTGCTGTTC	16	368
HAe6702PCR2	CCACAGGAGCGACCCAGAAA	16	116
HAe6702PCR1	GGTTTGTTGTATTGCTGTTC	16	368
HAe6703PCR2	CAGAGGAGGAGGATGAAATAGTA	16	656
HAe6703PCR1	GCACAACCGAAGCGTAGAGTCACAC	16	741
HAe6704PCR2	CAGAGGAGGAGGATGAAATAGA	16	656
HAe6704PCR1	GCACAACCGAAGCGTAGAGTCA	16	741
H18e6701PCR2	ACGATGAAATAGATGGAGTT	18	702
H18e6701PCR1	CACGGACACACAAAGGACAG	18	869
H18e6702PCR2	GAAAACGATGAAATAGATGGAG	18	698
H18e6702PCR1	ACACCACGGACACACAAAGGACAG	18	869
H18e6703PCR2	TTCCGGTTGACCTTCTATGT	18	651
H18e6703PCR1	GGTCGTCTGCTGAGCTTTCT	18	817
H18e6704PCR2	GCAAGACATAGAAATAACCTG	18	179
H18e6704PCR1	ACCCAGTGTTAGTTAGTT	18	379
H31e6701PCR2	GGAAATACCCTACGATGAAC	31	164
H31e6701PCR1	GGACACAACGGTCTTTGACA	31	423
H31e6702PCR2	GGAAATACCCTACGATGAACTA	31	164
H31e6702PCR1	CTGGACACAACGGTCTTTGACA	31	423
H31e6703PCR2	ACTGACCTCCACTGTTATGA	31	617
H31e6703PCR1	TATCTACTTGTGTGCTCTGT	31	766
H31e6704PCR2	TGACCTCCACTGTTATGAGCAATT	31	619
H31e6704PCR1	TGCGAATATCTACTTGTGTGCTCT GT	31	766
H31e6705PCR2	ACTGACCTCCACTGTTAT	31	617
H31e6705PCR1	CACGATTCCAAATGAGCCCAT	31	809
H33e6701PCR2	TATCCTGAACCAACTGACCTAT	33	618
H33e6701PCR1	TTGACACATAAACGAACTG	33	763
H33e6703PCR2	TCCTGAACCAACTGACCTAT	33	620
H33e6703PCR1	CCCATAAGTAGTTGCTGTAT	33	807
H33e6702PCR2	GACCTTTGTGTCCTCAAGAA	33	431
H33e6702PCR1	AGGTCAGTTGGTTCAGGATA	33	618
H35e6701PCR2	ATTACAGCGGAGTGAGGTAT	35	217
H35e6701PCR1	GTCTTTGCTTTTCAACTGGA	35	442
H35e6702PCR2	TCAGAGGAGGAGGAAGATACTA	35	655

Primer name	Sequence	HPV type	nt
H35e6702PCR1	GATTATGCTCTCTGTGAACA	35	844
H35e6703PCR2	CCCGAGGCAACTGACCTATA	35	610
H35e6703PCR1	GTCAATGTGTGTGCTCTGTA	35	770
H52e6701PCR2	TTGTGTGAGGTGCTGGAAGAAT	52	144
H52e6701PCR1	CCCTCTCTTCTAATGTTT	52	358
H52e6702PCR2	GTGCCTACGCTTTTTATCTA	52	296
H52e6702PCR1	GGGGTCTCCAACACTCTGAACA	52	507
H58e6701PCR2	TCAGGCGTTGGAGACATC	58	157
H58e6701PCR1	AGCAATCGTAAGCACACT	58	301
H58e6702PCR2	TCTGTGCATGAAATCGAA	58	173
H58e6702PCR1	AGCACACTTTACATACTG	58	291
HBe6701PCR2	TACACTGCTGGACAACAT	B(11)	514
HBe6701PCR1	TCATCTTCTGAGCTGTCT	B(11)	619
HBe6702PCR2	TACACTGCTGGACAACATGCA	B(11)	514
HBe6702PCR1	GTCACATCCACAGCAACAGGTCA	B(11)	693
HBe6703PCR2	TGACCTGTTGCTGTGGATGTGA	B(11)	693
HBe6703PCR1	TACCTGAATCGTCCGCCAT	B(11)	832
HCe6701PCR2	CATGCCATAAATGTATAGA	C (18 39 45	295
HCe6701PCR1	CACCGCAGGCACCTTATTAA	C (18 39 45	408
H39e6701PCR2	GCAGACGACCACTACAGCAAA	39	210
H39e6701PCR1	ACACCGAGTCCGAGTAATA	39	344
H39e6702PCR2	TATTACTCGGACTCGGTGT	39	344
H39e6702PCR1	CTTGGGTTTCTCTTCGTGTTA	39	558
H39e6703PCR2	GAAATAGATGAACCCGACCA	39	703
H39e6703PCR1	GCACACCACGGACACACAAA	39	886
H45e6701PCR2	AACCATTGAACCCAGCAGAAA	45	430
H45e6701PCR1	TCTTTCTTGCCGTGCCTGGTCA	45	527
H45e6702PCR2	GAAACCATTGAACCCAGCAGAAAA	45	428
H45e6702PCR1	TTGCTATACTTGTGTTTCCCTACG	45	558
H45e6703PCR2	GTTGACCTGTTGTGTTACCAGCAAT	45	656
H45e6703PCR1	CACCACGGACACACAAAGGACAAG	45	868
H45e6704PCR2	CTGTTGACCTGTTGTGTTACGA	45	654
H45e6704PCR1	CCACGGACACACAAAGGACAAG	45	868
H45e6705PCR2	GTTGACCTGTTGTGTTACGA	45	656
H45e6705PCR1	ACGGACACACAAAGGACAAG	45	868
H51e6701PCR2	GGAGGAGGATGAAGTAGATA	51	658
H51e6701PCR1	GCCCATTAACATCTGCTGTA	51	807
H51e6702PCR2	AGAGGAGGAGGATGAAGTAGATA	51	655
H51e6702PCR1	ACGGGCAAACCAGGCTTAGT	51	829
H56e6701PCR2	TTGGGGTGCTGGAGACAAACATCT	56	519
H56e6701PCR1	TTCATCCTCATCCTCATCCTCTGA	56	665
H56e6702PCR2	TGGGGTGCTGGAGACAAACATC	56	520
H56e6702PCR1	CATCCTCATCCTCATCCTCTGA	56	665
H56e6703PCR2	TTGGGGTGCTGGAGACAAACAT	56	519
H56e6703PCR1	CCACAAACTTACACTCACAACA	56	764
H56e6704PCR2	GATTTTCCTTATGCAGTGTG	56	279
H56e6704PCR1	GACATCTGTAGCACCTTATT	56	410

Primer-pairs and primer-probe sets

The invention further provides primer-pairs and primer/probe sets for use in the detection of HPV E6 transcripts.

A "primer-pair" is taken to mean two primers which may be used in combination for amplification of a portion of an HPV E6 transcript, for example by NASBA or RT-PCR. The individual oligonucleotide primers making up the primer-pair may be supplied separately, e.g. in separate containers. A primer-pair may also be supplied as a homogenous mixture of the two primers, this mixture may include additional reagents required for the amplification reaction, as discussed below.

A "primer/probe set" is taken to mean a set of oligonucleotides comprising a primer-pair, as defined above, and at least one oligonucleotide probe which is suitable for use in detection of an amplification product generated by use of the primer-pair. The individual oligonucleotides making up the primer/probe set may be supplied separately, e.g. in separate containers or as a homogenous mixture.

In this context "primer" is taken to encompass primers suitable for use in PCR and primers suitable for use in NASBA.

The term "probe" may encompass any of the probe types described herein, including molecular beacons probes suitable for use in real-time NASBA (see below) and capture probes for immobilisation of NASBA amplification products.

Specific primer-pairs provided by the invention are given below, together with suitable probes which may be used in the detection of amplification products

generated using the primer-pair. In preferred embodiments, the primer-pairs listed below may comprise a NASBA P1 primer and a NASBA P2 primer or two PCR primers. The most preferred specific primer combinations are listed, using the primer names given in Tables 2 and 3. However, it is not intended to limit the scope of the invention to these particular combinations:

Primer-pairs and probes for use in the detection of mRNA transcripts from the E6 gene of HPV 16:

(1) an oligonucleotide primer comprising sequence number 1 and an oligonucleotide primer comprising sequence number 2; oligonucleotide probe comprising sequence number 5.

Preferred NASBA primers: HAe6701p1 and HAe6701p2
Preferred PCR primers: HAe6701PCR1 and HAe6701PCR2

(2) an oligonucleotide primer comprising sequence number 3 and an oligonucleotide primer comprising sequence number 4; oligonucleotide probe comprising sequence number 6.

Preferred NASBA primers: HAe6702p1 and HAe6702p2
Preferred PCR primers: HAe 6702PCR1 and HAe6702PCR2

(3) an oligonucleotide primer comprising sequence number 7 and an oligonucleotide primer comprising sequence number 8; oligonucleotide probe comprising sequence number 9.

Preferred NASBA primers: HAe6703p1 and HAe6703p2
Preferred PCR primers: HAe6703PCR1 and HAe6703PCR2

(4) an oligonucleotide primer comprising sequence number 10 and an oligonucleotide primer comprising sequence number 11; oligonucleotide probe comprising sequence number 12.

Preferred NASBA primers: HAe6704p1 and HAe6704p2
Preferred PCR primers: HAe6704PCR1 and HAe6704PCR2

(5) an oligonucleotide primer comprising one of sequence numbers 126, 127, 128 or 129 and an oligonucleotide primer comprising sequence number 1 or sequence number 3.

(6) an oligonucleotide primer comprising sequence number 2 or sequence number 4 and an oligonucleotide primer comprising one of sequence numbers 130, 131, 132 or 133.

Primer-pairs for use in the detection of mRNA transcripts from the E6 gene of HPV 18:

(7) an oligonucleotide primer comprising sequence number 13 and an oligonucleotide primer comprising sequence number 14; oligonucleotide probe comprising sequence number 15.

Preferred NASBA primers: H18e6701p1 and H18e6701p2
Preferred PCR primers: H18e6701PCR1 and H18e6701PCR2

(8) an oligonucleotide primer comprising sequence number 16 and an oligonucleotide primer comprising sequence number 17; oligonucleotide probe comprising sequence number 18.

Preferred NASBA primers: H18e6702p1 and H18e6702p2
Preferred PCR primers: H18e6702PCR1 and H18e6702PCR2

(9) an oligonucleotide primer comprising sequence number 19 and an oligonucleotide primer comprising sequence number 20.

Preferred NASBA primers: H18e6703p1 and H18e6703p2
Preferred PCR primers: H1836703PCR1 and H18e6703PCR2

(10) an oligonucleotide primer comprising sequence number 21 and an oligonucleotide primer comprising sequence number 22; oligonucleotide probe comprising sequence number 23.

Preferred NASBA primers: H18e6704p1 and H18e6704p2
Preferred PCR primers: H18e6704PCR1 and H18e6704PCR2

Primer-pairs for use in the detection of mRNA transcripts from the E6 gene of HPV 31:

(11) an oligonucleotide primer comprising sequence number 24 and an oligonucleotide primer comprising sequence number 25; oligonucleotide probe comprising sequence number 26.

Preferred NASBA primers: H31e6701p1 and H31e6701p2
Preferred PCR primers: H31e6701PCR1 and H31e6701PCR2

(12) an oligonucleotide primer comprising sequence number 27 and an oligonucleotide primer comprising sequence number 28; oligonucleotide probe comprising sequence number 29.

Preferred NASBA primers: H31e6702p1 and H31e6702p2
Preferred PCR primers: H31e6702PCR1 and H3136702PCR2

(13) an oligonucleotide primer comprising sequence number 30 and an oligonucleotide primer comprising

sequence number 31; oligonucleotide probe comprising sequence number 32.

Preferred NASBA primers: H31e6703p1 and H31e6703p2
Preferred PCR primers: H31e6703PCR1 and H31e6703PCR2

(14) an oligonucleotide primer comprising sequence number 33 and an oligonucleotide primer comprising sequence number 34; oligonucleotide probe comprising sequence number 35.

Preferred NASBA primers: H31e6704p1 and H31e6704p2
Preferred PCR primers: H31e6704PCR1 and H312e6704PCR2

(15) an oligonucleotide primer comprising sequence number 36 and an oligonucleotide primer comprising sequence number 37;

Preferred NASBA primers: H31e6705p1 and H31e6705p2
Preferred PCR primers: H31e6705PCR1 and H31e6705PCR2

Primer-pairs for use in the detection of mRNA transcripts from the E6 gene of HPV 33:

(16) an oligonucleotide primer comprising sequence number 38 and an oligonucleotide primer comprising sequence number 39; oligonucleotide probe comprising sequence number 40.

Preferred NASBA primers: H33e6701p1 and H33e6701p2
Preferred PCR primers: H33e6701PCR1 and H33e6701PCR2

(17) an oligonucleotide primer comprising sequence number 41 and an oligonucleotide primer comprising sequence number 42; oligonucleotide probe comprising sequence number 43.

Preferred NASBA primers: H33e6703p1 and H33e6703p2
Preferred PCR primers: H33e6703PCR1 and H33e6703PCR2

(18) an oligonucleotide primer comprising sequence number 44 and an oligonucleotide primer comprising sequence number 45; oligonucleotide probe comprising sequence number 46.

Preferred NASBA primers: H33e6702p1 and H33e6702p2
Preferred PCR primers: H33e6702PCR1 and H33e6702PCR2

Primer-pairs for use in the detection of mRNA transcripts from the E6 gene of HPV 35:

(19) an oligonucleotide primer comprising sequence number 47 and an oligonucleotide primer comprising sequence number 48; oligonucleotide probe comprising sequence number 53.

Preferred NASBA primers: H35e6701p1 and H35e6701p2
Preferred PCR primers: H35e6701PCR1 and H35e6701PCR2

(20) an oligonucleotide primer comprising sequence number 49 and an oligonucleotide primer comprising sequence number 50; oligonucleotide probe comprising sequence number 54.

Preferred NASBA primers: H35e6702p1 and H35e6702p2
Preferred PCR primers: H35e6702PCR1 and H35e6702PCR2

(21) an oligonucleotide primer comprising sequence number 51 and an oligonucleotide primer comprising sequence number 52; oligonucleotide probe comprising sequence number 55.

Preferred NASBA primers: H35e6703p1 and H35e6703p2
Preferred PCR primers: H35e6703PCR1 and H35e6703PCR2

Primer-pairs for use in the detection of mRNA transcripts from the E6 gene of HPV 52:

(22) an oligonucleotide primer comprising sequence number 56 and an oligonucleotide primer comprising sequence number 57; oligonucleotide probe comprising sequence number 58.

Preferred NASBA primers: H52e6701p1 and H52e6701p2
Preferred PCR primers: H52e6701PCR1 and H52e6701PCR2

(23) an oligonucleotide primer comprising sequence number 59 and an oligonucleotide primer comprising sequence number 60; oligonucleotide probe comprising sequence number 61.

Preferred NASBA primers: H52e6702p1 and H52e6702p2
Preferred PCR primers: H52e6702PCR1 and H52e6702PCR2

Primer-pairs for use in the detection of mRNA transcripts from the E6 gene of HPV 58:

(24) an oligonucleotide primer comprising sequence number 62 and an oligonucleotide primer comprising sequence number 63; oligonucleotide probe comprising sequence number 66.

Preferred NASBA primers: H58e6701p1 and H58e6701p2
Preferred PCR primers: H58e6701PCR1 and H58e6701PCR2

(25) an oligonucleotide primer comprising sequence number 64 and an oligonucleotide primer comprising sequence number 65; oligonucleotide probe comprising sequence number 67.

Preferred NASBA primers: H58e6702p1 and H58e6702p2
Preferred PCR primers: H58e6702PCR1 and H58e6702PCR2

Primer-pairs for use in the detection of mRNA transcripts from the E6 gene of HPV 51:

(26) an oligonucleotide primer comprising sequence number 104 and an oligonucleotide primer comprising sequence number 105; oligonucleotide probe comprising sequence number 108.

Preferred NASBA primers: H51e6701p1 and H51e6701p2
Preferred PCR primers: H51e6701PCR1 and H51e6701PCR2

(27) an oligonucleotide primer comprising sequence number 106 and an oligonucleotide primer comprising sequence number 107; oligonucleotide probe comprising sequence number 109.

Preferred NASBA primers: H51e6702p1 and H51e6702p2
Preferred PCR primers: H51e6702PCR1 and H51e6702PCR2

Primer-pairs for use in the detection of mRNA transcripts from the E6 gene of HPV 56:

(28) an oligonucleotide primer comprising sequence number 110 and an oligonucleotide primer comprising sequence number 111; oligonucleotide probe comprising sequence number 116.

Preferred NASBA primers: H56e6701p1 and H56e6701p2
Preferred PCR primers: H56e6701PCR1 and H56e6701PCR2

(29) an oligonucleotide primer comprising sequence number 112 and an oligonucleotide primer comprising sequence number 113; oligonucleotide probe comprising sequence number 117.

Preferred NASBA primers: H56e6702p1 and H56e6702p2
Preferred PCR primers: H56e6702PCR1 and H56e6702PCR2

(30) an oligonucleotide primer comprising sequence number 114 and an oligonucleotide primer comprising sequence number 115; oligonucleotide probe comprising sequence number 118 or sequence number 119.

Preferred NASBA primers: H56e6703p1 and H56e6703p2
Preferred PCR primers: H56e6703PCR1 and H56e6703PCR2

(31) an oligonucleotide primer comprising sequence number 120 and an oligonucleotide primer comprising sequence number 121; oligonucleotide probe comprising sequence number 122.

Preferred NASBA primers: H56e6704p1 and H56e6704p2
Preferred PCR primers: H56e6704PCR1 and H56e6704PCR2

Primer-pairs for use in the detection of mRNA transcripts from the E6 gene of HPV 39:

(32) an oligonucleotide primer comprising sequence number 80 and an oligonucleotide primer comprising sequence number 81; oligonucleotide probe comprising sequence number 82.

Preferred NASBA primers: H39e6701p1 and H39e6701p2
Preferred PCR primers: H39e6701PCR1 and H39e6701PCR2

(33) an oligonucleotide primer comprising sequence number 83 and an oligonucleotide primer comprising sequence number 84; oligonucleotide probe comprising sequence number 85.

Preferred NASBA primers: H39e6702p1 and H39e6702p2
Preferred PCR primers: H39e6702PCR1 and H39e6702PCR2

(34) an oligonucleotide primer comprising sequence number 86 and an oligonucleotide primer comprising

sequence number 87; oligonucleotide probe comprising sequence number 88.

Preferred NASBA primers: H39e6703p1 and H39e6703p2
Preferred PCR primers: H39e6703PCR1 and H39e6703PCR2

Primer-pairs for use in the detection of mRNA transcripts from the E6 gene of HPV 45:

(35) an oligonucleotide primer comprising sequence number 89 and an oligonucleotide primer comprising sequence number 90; oligonucleotide probe comprising sequence number 93.

Preferred NASBA primers: H45e6701p1 and H45e6701p2
Preferred PCR primers: H45e6701PCR1 and H45e6701PCR2

(36) an oligonucleotide primer comprising sequence number 91 and an oligonucleotide primer comprising sequence number 92; oligonucleotide probe comprising sequence number 94.

Preferred NASBA primers: H45e6702p1 and H45e6702p2
Preferred PCR primers: H45e6702PCR1 and H45e6702PCR2

(37) an oligonucleotide primer comprising sequence number 95 and an oligonucleotide primer comprising sequence number 96; oligonucleotide probe comprising sequence number 101.

Preferred NASBA primers: H45e6703p1 and H45e6703p2
Preferred PCR primers: H45e6703PCR1 and H45e6703PCR2

(38) an oligonucleotide primer comprising sequence number 97 and an oligonucleotide primer comprising sequence number 98; oligonucleotide probe comprising sequence number 102.

Preferred NASBA primers: H45e6704p1 and H45e6704p2
Preferred PCR primers: H45e6704PCR1 and H45e6704PCR2

(39) an oligonucleotide primer comprising sequence number 99 and an oligonucleotide primer comprising sequence number 100; oligonucleotide probe comprising sequence number 103.

Preferred NASBA primers: H45e6705p1 and H45e6705p2
Preferred PCR primers: H45e6705PCR1 and H45e6705PCR2

Primer-pairs for use in the detection of mRNA transcripts from the E6 gene of group B HPV:

(40) an oligonucleotide primer comprising sequence number 68 and an oligonucleotide primer comprising sequence number 69; oligonucleotide probe comprising sequence number 72.

Preferred NASBA primers: HBe6701p1 and HBe6701p2
Preferred PCR primers: HBe6701PCR1 and HBe6701PCR2

(41) an oligonucleotide primer comprising sequence number 70 and an oligonucleotide primer comprising sequence number 71; oligonucleotide probe comprising sequence number 73.

Preferred NASBA primers: HBe6702p1 and HBe6702p2
Preferred PCR primers: HBe6702PCR1 and HBe6702PCR2

(42) an oligonucleotide primer comprising sequence number 74 and an oligonucleotide primer comprising sequence number 75; oligonucleotide probe comprising sequence number 76.

Preferred NASBA primers: HBe6703p1 and HBe6703p2
Preferred PCR primers: HBe6703PCR1 and HBe6703PCR2

Primer-pair for use in the detection of mRNA transcripts from the E6 gene of group C HPV:

(43) an oligonucleotide primer comprising sequence number 77 and an oligonucleotide primer comprising sequence number 78; oligonucleotide probe comprising sequence number 79.

Preferred NASBA primers: HCe6701p1 and HCe6701p2
Preferred PCR primers: HCe6701PCR1 and HCe6701PCR2

Methods of detecting HPV

In a further aspect the invention provides a method for detecting HPV mRNA in a test sample suspected of containing HPV which comprises performing an amplification reaction on the test sample to amplify a portion of the mRNA transcribed from the E6 gene of HPV, wherein the amplification reaction is performed using one of the primer-pairs provided by the invention, as defined above.

Preferred amplification techniques which may be used to amplify a portion of the E6 mRNA are RT-PCR or NASBA.

The "test sample suspected of containing HPV" will most commonly be a clinical sample, for example a cervical scraping in the cervical screening field. The amplification reaction will preferably be carried out on a preparation of nucleic acid isolated from the test sample. The preparation of nucleic acid must include mRNA, however it need not be a preparation of purified poly A+ mRNA and preparations of total RNA or crude preparations of total nucleic acid containing both RNA and genomic DNA are also suitable as starting material for a NASBA reaction. Essentially any technique known in the art for the isolation of a

preparation of nucleic acid including mRNA may be used to isolate nucleic acid from the test sample. A preferred technique is the "Boom" isolation method described in US-A-5,234,809 and EP-B-0389,063. This method, which can be used to isolate a nucleic acid preparation containing both RNA and DNA, is based on the nucleic acid binding properties of silicon dioxide particles in the presence of the chaotropic agent guanidine thiocyanate (GuSCN).

Methods for the detection of HPV in a test sample using the NASBA technique will generally comprise the following steps:

(a) assembling a reaction medium comprising a primer-pair according to the invention, an RNA directed DNA polymerase, a ribonuclease that hydrolyses the RNA strand of an RNA-DNA hybrid without hydrolysing single or double stranded RNA or DNA, an RNA polymerase that recognises said promoter, and ribonucleoside and deoxyribonucleoside triphosphates;

(b) incubating said reaction medium with a preparation of nucleic acid isolated from a test sample suspected of containing HPV under reaction conditions which permit a NASBA amplification reaction; and

(c) detecting and/or quantitatively measuring any HPV-specific product of the NASBA amplification reaction.

Detection of the specific product(s) of the NASBA reaction (i.e. sense and/or antisense copies of the target RNA) may be carried out in a number of different ways. In one approach the NASBA product(s) may be detected with the use of an HPV-specific hybridisation probe capable of specifically annealing to the NASBA product. The hybridisation probe may be attached to a revealing label, for example a

fluorescent, luminescent, radioactive or chemiluminescent compound or an enzyme label or any other type of label known to those of ordinary skill in the art. The precise nature of the label is not critical, but it should be capable of producing a signal detectable by external means, either by itself or in conjunction with one or more additional substances (e.g. the substrate for an enzyme).

Also within the scope of the invention is so-called "real-time NASBA" which allows continuous monitoring of the formation of the product of the NASBA reaction over the course of the reaction. In a preferred embodiment this may be achieved using a "molecular beacons" probe comprising an HPV-specific sequence capable of annealing to the NASBA product, a stem-duplex forming oligonucleotide sequence and a pair of fluorescer/quencher moieties, as known in the art described herein. If the molecular beacons probe is added to the reaction mixture prior to amplification it may be possible to monitor the formation of the NASBA product in real-time (Leone *et al.*, Nucleic Acids Research, 1998, Vol 26, 2150-2155).

In a further approach, the molecular beacons technology may be incorporated into the primer 2 oligonucleotide allowing real-time monitoring of the NASBA reaction without the need for a separate hybridisation probe.

In a still further approach the products of the NASBA reaction may be monitored using a generic labelled detection probe which hybridises to a nucleotide sequence in the 5' terminus of the primer 2 oligonucleotide. This is equivalent to the "NucliSens™" detection system supplied by Organon Teknika. In this system specificity for NASBA

products derived from the target HPV mRNA may be conferred by using HPV-specific capture probes comprising probe oligonucleotides as described herein attached to a solid support such as a magnetic microbead. Most preferably the generic labelled detection probe is the ECL™ detection probe supplied by Organon Teknika. NASBA amplicons are hybridized to the HPV-specific capture probes and the generic ECL probe (via a complementary sequence on primer 2). Following hybridization the bead/amplicon/ECL probe complexes may be captured at the magnet electrode of an automatic ECL reader (e.g. the NucliSens™ reader supplied by Organon Teknika. Subsequently, a voltage pulse triggers the ECL™ reaction.

Also provided by the invention are reagent kits for use in the detection of HPV by NASBA, the kits comprising a primer-pair cocktail according to the invention. The reagent kits may further comprise a mixture of enzymes required for the NASBA reaction, specifically an enzyme mixture containing an RNA directed DNA polymerase (e.g. a reverse transcriptase), a ribonuclease that hydrolyses the RNA strand of an RNA-DNA hybrid without hydrolysing single or double stranded RNA or DNA (e.g. RNaseH) and an RNA polymerase. The RNA polymerase should be one which recognises the promoter sequence present in the 5' terminal region of the NASBA P1 primer oligonucleotides in the oligonucleotide primer sets supplied in the reagent kit. The kit may also comprise a supply of NASBA buffer containing the ribonucleosides and deoxyribonucleosides required for RNA and DNA synthesis. The composition of a standard NASBA reaction buffer will be well known to those skilled in the art.

In certain embodiments the kit may further contain one or more capture probes, comprising a probe oligonucleotide attached to a solid support as described above, for immobilising the products of a specific NASBA reaction. The kit may still further contain labelled generic detection probes. Advantageously, the detection probes may comprise a sequence of nucleotides complementary to a non-HPV sequence present at the 5' terminal end of the NASBA P2 primer oligonucleotides present in the reagent kit.

In still further embodiments the kit may further contain one or more molecular beacon probes according to the invention. The molecular beacon probes may be supplied as a separate reagent within the kit. Alternatively, the NASBA primers and molecular beacons probe may be supplied as a primer/probe mixture. Such a mixture including the NASBA P1 and P2 primers and also a molecular beacons probe is convenient for use in "real-time" NASBA, wherein the NASBA amplification reaction and detection of an amplification product are performed simultaneously in a single reaction vessel.

The invention will be further understood with reference to the following, non-limiting, Example:

<u>Example 1-Real-time NASBA</u>

<u>Collection and preparation of clinical samples</u>

Cervical cytobrush samples are collected in 9 ml lysis buffer (5M Guanidine thiocyanate) prior to RNA/DNA extraction. Since RNA is best protected in the 5M guanidine thiocyanate at -70°C only 1 ml of the total volume of sample is used for each extraction round. 2-3 tubes with the RNA/DNA are stored at -167°C and the rest stored at -70°C.

RNA and DNA were automatically isolated according to the "Booms" isolation method from Organon Teknika (Organon Teknika B.V., Boselind 15, P.O. Box 84, 5280 AB Baxtel, The Netherlands; now Biomérieux, 69280 Marcy l'Etoile, France).

The following procedure was carried out using reagents from the Nuclisens™ Basic Kit, supplied by Organon Teknika. Procedure for n=10 samples:-

1. Prepare enzyme solution.
Add 55 µl of enzyme diluent (from Nuclisens™ Basic Kit; contains sorbitol in aqueous solution) to each of 3 lyophilized enzyme spheres (from Nuclisens™ Basic Kit; contains AMV-RT, RNase H, T7 RNA polymerase and BSA). Leave this enzyme solution at least for 20 minutes at room temperature. Gather the enzyme solutions in one tube, mix well by flicking the tube with your finger, spin down briefly and use within 1 hour. Final concentrations in the enzyme mix are 375 mM sorbitol, 2.5 µg BSA, 0.08 U RNase H, 32 U T7 RNA polymerase and 6.4 U AMV-reverse transcriptase.

2. Prepare reagent sphere/KCl solution.
For 10 samples: add 80 µl reagent sphere diluent (from Nuclisens™ Basic Kit; contains Tris/HCl (pH 8.5), 45% DMSO) to the lyophilized reagent sphere (from Nuclisens™ Basic Kit; contains nucleotides, dithiotreitol and $MgCl_2$) and immediately vortex well. Do this with 3 reagent spheres and mix the solutions in one tube.

Add 3 µl NASBA water (from Nuclisens™ Basic Kit) to the reconstituted reagent sphere solution and mix well.

Add 56 μl of KCl stock solution (from Nuclisens™ Basic Kit) and mix well. Use of this KCl/water mixture will result in NASBA reactions with a final KCl concentration of 70 mM. Final concentrations in the reagent/KCl solution are 1 mM of each dNTP, 2 mM of ATP, UTP and CTP, 1.5 mM GTP, and 0.5 mM ITP, 0.5 mM dithiotreitol, 70 mM KCl, 12 mM $MgCl_2$, 40 mM Tris-HCl (pH 8.5).

3. Prepare primer/probe solution containing target-specific primers and molecular beacon probe.
For each target reaction transfer 91 μl of the reagent sphere/KCl solution (prepared in step 2) into a fresh tube. Add 25 μl of primers/molecular beacon probe solution (to give final concentration of ~0.1-0.5 μM each of the sense and antisense primers and ~ 15-70 pmol molecular beacon probe per reaction). Mix well by vortexing. Do not centrifuge.

In case less than 10 target RNA amplifications are being performed refer to the table below for the appropriate amounts of reagent sphere solution, KCl/water solution and primers to be used. Primer solutions should be used within 30 minutes after preparation.

Reactions (n)	Reagent sphere solution (μl)	KCl/water (μl)	Primer mix (μl)
10	80	30	10
9	72	27	9
8	64	24	8
7	56	21	7
6	48	18	6
5	40	15	5
4	32	12	4
3	24	9	3
2	16	6	2
1	8	3	1

4. Addition of samples

For each target RNA reaction:

In a 96 well microtiter plate pipette 10 μl of the primer/probe solution (prepared in step 3) into each of 10 wells. Add 5 μl nucleic acid extract to each well. Incubate the microtiter plate for 4 minutes at 65 ± 1 °C. Cool to at 41 ± 0.5 °C for 4 minutes. Then to each well add 5 μl enzyme solution. Immediately place the microtiter plate in a fluorescent detection instrument (e.g. NucliSens™ EasyQ Analyzer) and start the amplification.

Claims:

1. An oligonucleotide molecule for use in the detection of mRNA transcribed from the E6 gene of a human papillomavirus, the oligonucleotide comprising any one of sequence numbers 1-133.

2. An oligonucleotide primer for use in the detection of mRNA transcribed from the E6 gene of a human papillomavirus, the oligonucleotide primer being selected from:

(i) a NASBA P1 primer comprising one of sequence numbers 2, 4, 8, 11, 14, 17, 20, 22, 25, 28, 31, 34, 37, 39, 42, 45, 48, 50, 52, 57, 60, 63, 65, 69, 71, 75, 78, 81, 84 87, 90, 92, 96, 98, 100, 105, 107, 111, 113, 115, 121, 126, 127, 128 Or 129;

(ii) a NASBA P2 primer comprising one of sequence numbers 1, 3, 7, 10, 13, 16, 19, 21, 24, 27, 30, 33, 36, 38, 41, 44, 47, 49, 51, 56, 59, 62, 64, 68, 70, 74, 77, 80, 83, 86, 89, 91, 95, 97, 99, 104, 106, 110, 112, 114, 120, 103, 131, 132 or 133;

(iii) a PCR primer comprising one of sequence numbers 1, 3, 7, 10, 13, 16, 19, 21, 24, 27, 30, 33, 36, 38, 41, 44, 47, 49, 51, 56, 59, 62, 64, 68, 70, 74, 77, 80, 83, 86, 89, 91, 95, 97, 99, 104, 106, 110, 112, 114, 120, 2, 4, 8, 11, 14, 17, 20, 22, 25, 28, 31, 34, 37, 39, 42, 45, 48, 50, 52, 57, 60, 63, 65, 69, 71, 75, 78, 81, 84 87, 90, 92, 96, 98, 100, 105, 107, 111, 113, 115, 121, 126, 127, 128, 129, 130, 131, 132 or 133.

3. An oligonucleotide primer according to claim 2 which is a NASBA P1 primer having the sequence AATTCTAATACGACTCACTATAGGGAGAAGG-SEQ, wherein SEQ represents any one of sequence numbers 2, 4, 8, 11,

14, 17, 20, 22, 25, 28, 31, 34, 37, 39, 42, 45, 48, 50, 52, 57, 60, 63, 65, 69, 71, 75, 78, 81, 84 87, 90, 92, 96, 98, 100, 105, 107, 111, 113, 115, 121, 126, 127, 128 or 129.

4. An oligonucleotide primer according to claim 2 which is a NASBA P2 primer having the sequence GATGCAAGGTCGCATATGAG-SEQ wherein SEQ represents any one of sequence numbers 1, 3, 7, 10, 13, 16, 19, 21, 24, 27, 30, 33, 36, 38, 41, 44, 47, 49, 51, 56, 59, 62, 64, 68, 70, 74, 77, 80, 83, 86, 89, 91, 95, 97, 99, 104, 106, 110, 112, 114, 120, 130, 131, 132 or 133.

5. An oligonucleotide probe for use in the detection of mRNA transcribed from the E6 gene of a human papillomavirus comprising one of sequence numbers: 5, 6, 9, 12, 15, 18, 23, 26, 29, 32, 35, 40, 43, 46, 53, 54, 55, 58, 61, 66, 67, 72, 73, 76, 82, 85, 88, 93, 94, 101, 102, 103, 108, 109, 116, 117, 118, 119, 122, 130, 131, 132 or 133.

6. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of HPV 16, comprising one of the following combinations:

an oligonucleotide primer comprising sequence number 1 and an oligonucleotide primer comprising sequence number 2;
an oligonucleotide primer comprising sequence number 3 and an oligonucleotide primer comprising sequence number 4;
an oligonucleotide primer comprising sequence number 7 and an oligonucleotide primer comprising sequence number 8;

an oligonucleotide primer comprising sequence number 10 and an oligonucleotide primer comprising sequence number 11;
an oligonucleotide primer comprising one of sequence numbers 126, 127, 128 or 129 and an oligonucleotide primer comprising sequence number 1 or sequence number 3; or
an oligonucleotide primer comprising sequence number 2 or sequence number 4 and an oligonucleotide primer comprising one of sequence numbers 130, 131, 132 or 133.

7. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of HPV 18, comprising one of the following combinations:

an oligonucleotide primer comprising sequence number 13 and an oligonucleotide primer comprising sequence number 14;
an oligonucleotide primer comprising sequence number 16 and an oligonucleotide primer comprising sequence number 17;
an oligonucleotide primer comprising sequence number 19 and an oligonucleotide primer comprising sequence number 20; or
an oligonucleotide primer comprising sequence number 21 and an oligonucleotide primer comprising sequence number 22.

8. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of HPV 31, comprising one of the following combinations:

an oligonucleotide primer comprising sequence number 24 and an oligonucleotide primer comprising sequence number 25;

an oligonucleotide primer comprising sequence number 27 and an oligonucleotide primer comprising sequence number 28;
an oligonucleotide primer comprising sequence number 30 and an oligonucleotide primer comprising sequence number 31;
an oligonucleotide primer comprising sequence number 33 and an oligonucleotide primer comprising sequence number 34; or
an oligonucleotide primer comprising sequence number 36 and an oligonucleotide primer comprising sequence number 37.

9. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of HPV 33, comprising one of the following combinations:

an oligonucleotide primer comprising sequence number 38 and an oligonucleotide primer comprising sequence number 39;
an oligonucleotide primer comprising sequence number 41 and an oligonucleotide primer comprising sequence number 42; or
an oligonucleotide primer comprising sequence number 44 and an oligonucleotide primer comprising sequence number 45.

10. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of HPV 35, comprising one of the following combinations:

an oligonucleotide primer comprising sequence number 47 and an oligonucleotide primer comprising sequence number 48;
an oligonucleotide primer comprising sequence number 49 and an oligonucleotide primer comprising sequence number 50; or

an oligonucleotide primer comprising sequence number 51 and an oligonucleotide primer comprising sequence number 52.

11. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of HPV 52, comprising one of the following combinations:

an oligonucleotide primer comprising sequence number 56 and an oligonucleotide primer comprising sequence number 57; or
an oligonucleotide primer comprising sequence number 59 and an oligonucleotide primer comprising sequence number 60.

12. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of HPV 58, comprising one of the following combinations:

an oligonucleotide primer comprising sequence number 62 and an oligonucleotide primer comprising sequence number 63;
an oligonucleotide primer comprising sequence number 64 and an oligonucleotide primer comprising sequence number 65.

13. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of HPV 51, comprising one of the following combinations:

an oligonucleotide primer comprising sequence number 104 and an oligonucleotide primer comprising sequence number 105; or
an oligonucleotide primer comprising sequence number 106 and an oligonucleotide primer comprising sequence number 107.

14. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of HPV 56, comprising one of the following combinations:

an oligonucleotide primer comprising sequence number 110 and an oligonucleotide primer comprising sequence number 111;
an oligonucleotide primer comprising sequence number 112 and an oligonucleotide primer comprising sequence number 113;
an oligonucleotide primer comprising sequence number 114 and an oligonucleotide primer comprising sequence number 115;
an oligonucleotide primer comprising sequence number 120 and an oligonucleotide primer comprising sequence number 121.

15. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of HPV 39, comprising one of the following combinations:

an oligonucleotide primer comprising sequence number 80 and an oligonucleotide primer comprising sequence number 81;
an oligonucleotide primer comprising sequence number 83 and an oligonucleotide primer comprising sequence number 84; or
an oligonucleotide primer comprising sequence number 86 and an oligonucleotide primer comprising sequence number 87.

16. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of HPV 45, comprising one of the following combinations:

an oligonucleotide primer comprising sequence number 89 and an oligonucleotide primer comprising sequence number 90;
an oligonucleotide primer comprising sequence number 91 and an oligonucleotide primer comprising sequence number 92;
an oligonucleotide primer comprising sequence number 95 and an oligonucleotide primer comprising sequence number 96;
an oligonucleotide primer comprising sequence number 97 and an oligonucleotide primer comprising sequence number 98; or
an oligonucleotide primer comprising sequence number 99 and an oligonucleotide primer comprising sequence number 100.

17. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of group B HPV, comprising one of the following combinations:

an oligonucleotide primer comprising sequence number 68 and an oligonucleotide primer comprising sequence number 69;
an oligonucleotide primer comprising sequence number 70 and an oligonucleotide primer comprising sequence number 71; or
an oligonucleotide primer comprising sequence number 74 and an oligonucleotide primer comprising sequence number 75.

18. An oligonucleotide primer-pair for use in the detection of mRNA transcripts from the E6 gene of group C HPV, comprising the following combination:

an oligonucleotide primer comprising sequence number 77 and an oligonucleotide primer comprising sequence number 78.

19. An oligonucleotide primer-pair according to any one of claims 6 to 18 which comprises a NASBA P1 primer and a NASBA P2 primer.

20. A primer-pair according to claim 19 wherein the NASBA P1 primer includes the sequence AATTCTAATACGACTCACTATAGGGAGAAGG at the 5' end.

21. A primer/probe set comprising a primer-pair according to any one of claims 6 to 20 and at least one oligonucleotide probe specific for amplification products generated using the primer-pair.

22. A method of detecting HPV mRNA in a test sample suspected of containing HPV which comprises performing an amplification reaction on a preparation of nucleic acid isolated from the test sample to amplify a portion of the mRNA transcribed from the E6 gene of HPV, wherein the amplification reaction is performed using a primer-pair according to any one of claims 6 to 18.

23. A method according to claim 22 which comprises performing RT-PCR to amplify a portion of the mRNA transcribed from the E6 gene of HPV.

24. A method according to claim 126 which comprises performing NASBA to amplify a portion of the mRNA transcribed from the E6 gene of HPV .

25. A method according to claim 24 which comprises:

(a) assembling a reaction mixture comprising a primer set as defined in any one of claims 6 to 18, an RNA directed DNA polymerase, a ribonuclease that hydrolyses the RNA strand of an RNA-DNA hybrid without hydrolysing single or double stranded RNA or DNA, an RNA polymerase that recognises said promoter, and ribonucleoside and deoxyribonucleoside triphosphates;

(b) incubating said reaction mixture with a preparation of nucleic acid isolated from a test sample suspected of containing HPV under reaction conditions which permit a NASBA amplification reaction; and

(c) detecting and/or quantitatively measuring any HPV-specific product of the NASBA amplification reaction.

26. A method according to claim 25 wherein step (c) comprises real-time detection of an HPV-specific product of the NASBA amplification reaction.

27. A method according to claim 25 or claim 26 wherein the reaction mixture further comprises a molecular beacons probe oligonucleotide and the formation of any HPV-specific NASBA product in the NASBA reaction is monitored by detecting fluorescence from the fluorescent moiety included in the molecular beacons probe.

28. A method according to claim 25 or claim 26 which comprises the further step of capturing the NASBA reaction product by hybridisation to a probe oligonucleotide attached to a solid support.

29. A reagent kit for use in the detection of HPV by NASBA, the kit comprising an oligonucleotide primer-pair as defined in claim 19 and optionally an enzyme mixture comprising an RNA directed DNA

polymerase, a ribonuclease that hydrolyses the RNA strand of an RNA-DNA hybrid without hydrolysing single or double stranded RNA or DNA, and an RNA polymerase that recognises the promoter sequence present in at least one NASBA P1 primer oligonucleotide included in the reagent kit.

US006846910B2

(12) **United States Patent**
Zimmermann et al.

(10) **Patent No.: US 6,846,910 B2**
(45) **Date of Patent: Jan. 25, 2005**

(54) **ISOLATED PROTEINS CONTAINING PORTIONS OF FAPα AND OTHER PROTEINS**

(75) Inventors: **Rainer Zimmermann**, Mittelbiberach (DE); **John E. Park**, Biberach (DE); **Wolfgang Rettig**, Biberach (DE); **Lloyd J. Old**, New York, NY (US)

(73) Assignees: **Ludwig Institute for Cancer Research**, New York, NY (US); **Boehringer Ingelheim International GmbH**, Ingelheim em Rhein (DE)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 108 days.

(21) Appl. No.: **09/265,606**

(22) Filed: **Mar. 10, 1999**

(65) **Prior Publication Data**

US 2002/0034789 A1 Mar. 21, 2002

Related U.S. Application Data

(60) Division of application No. 08/940,391, filed on Oct. 1, 1997, now Pat. No. 5,965,373, which is a continuation-in-part of application No. 08/230,491, filed on Apr. 20, 1994, now Pat. No. 5,587,299.

(51) **Int. Cl.**[7] **C07K 14/745**; C07K 14/52; C07K 19/00

(52) **U.S. Cl.** **530/381**; 530/351; 424/192.1

(58) **Field of Search** 530/381, 351; 424/192.1

(56) **References Cited**

PUBLICATIONS

Niedermeyer et al. Mouse fibroblast–activation protein. European Journal of Biochemistry (Jun. 1998) vol. 254 (2), pp. 650–654.*

Ogata et al. Identification of the ative site residues in dipeptidyl peptidase IV by affinity labeling and site–directed mutagenesis. Biochemistry, vol. 31, pp. 2582–2587 (1992).*

Geysen et al. Cognitive features of continuous antigenic determinants. (J. of Molecular Recognition, vol. 1, pp. 32–40, 1988).*

Rettig et al. Cancer Research, vol. 53, pp. 3327–3335, 1993.*

* cited by examiner

Primary Examiner—Marjorie Moran
(74) *Attorney, Agent, or Firm*—Fulbright & Jaworski L.L.P.

(57) **ABSTRACT**

The invention involves dimeric forms of the protein known as fibroblast activation protein alpha, or "FAPα" and its uses.

5 Claims, 3 Drawing Sheets

FIG. 1

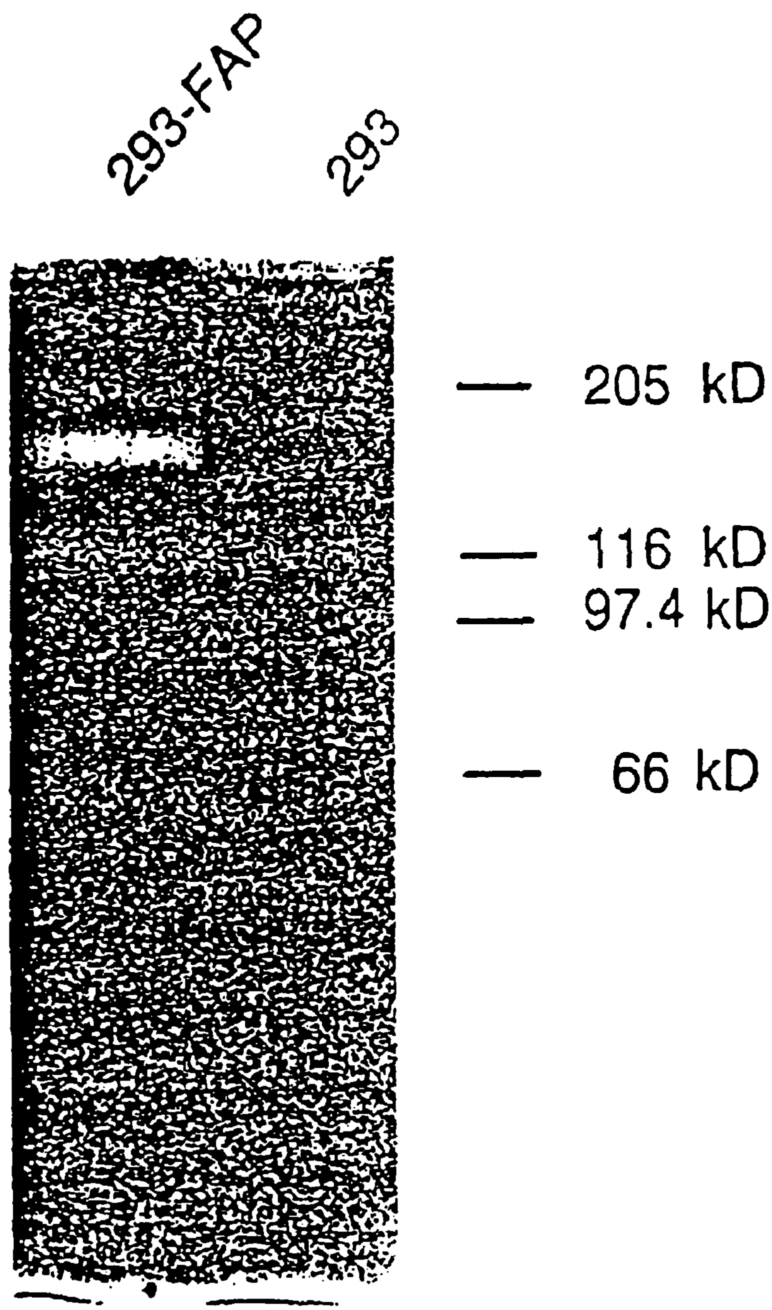
```
FAP    1 MKTWVKIVFGV*ATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILN  49
CD26   1 ---PW-VLL-LLGAA-LVTIITVPV--LNKGTDDATADSRKTY--T-Y-K  50

FAP   50 GTFSYKTFFPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSV*  98
CD26  51 N-YRL-LYSLR---DH---YKQ*E---LVF-A-Y-N-SVF-E-S-FDEFG  99

FAP   99 *NASNYGLSPDRQFVYLESDYSKLWRYSYTATYYIYDLSNGEFVRGNELP 147
CD26 100 HSIND-SI---G--IL--YN-V-Q--H----S-D----NKRQLITEERI- 149
                                                 fap-1
FAP  148 RPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGIP 197
CD26 150 NNT-WVT-----H-----WN-D--V-IE-NL-SYR--WT-K-DI-Y---T 199
                              fap-2
FAP  198 DWVYEEEMLPTKYALWWSPNGKFLAYAEFNDKDIPVIAYSYYGDE**QYP 245
CD26 200 -------VFSAYS--------T-----Q---TEV-L-E--F-S--SL--- 249

FAP  246 RTINIPYPKAGAKNPVVRIFIIDT***TYPAYVGPQEVPVPAMIASSDYY 292
CD26 250 K-VRV-------V--T-KF-VVN-DSLSSVTNATSIQITA--SMLIG-H- 299

FAP  293 FSWLTWVTDERVCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEES 342
CD26 300 LCDV--A-Q--IS----R-I--Y--MD---YD-SSGR-N-LVARQ---M- 349

FAP  343 RTGWAGGFFVSRPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITS 392
CD26 350 T---V-R-RP-E-H-TL-GN-F---I-NEE--R--C-FQIDKKDCTF--K 399

FAP  393 GKWEAINIFRVTQDSLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCH 442
CD26 400 -T--V-G-EAL-S-Y-Y-I---YKGM--G--L-K-QLSD-T*KVT-LS-E 448

FAP  443 LRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGRTDQEIKILEENK 492
CD26 449 -NP------SV---KE----Q-R-S---L-LY---SSVN-KGLRV--D-S 498
                     fap-3
FAP  493 ELENALKNIQLPKEEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQVYG 542
CD26 499 A-DKM-Q-V-M-SKKLDFIILN-TKF--Q-----H--K-------LD--A 548

FAP  543 GPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKL 592
CD26 549 -----KADT--RL--AT----T-NIIV-SF----SGY----IMH-IN-R- 598

FAP  593 GVYEVEDQITAVRKFIEMGFIDEKRIAIWGWSYEIRFITGPCIWNWSFQM 642
CD26 599 -TF------E-A-Q-SK---V-N----------GGYVTSMVLGSGSVGFK 648

FAP  643 WYSSGSSLQLGILRVCLHRE*IHGSPNKDDNLEHYKNSTVMARAEYFRNV 691
CD26 649 CGIAVAPVSRWEYYDSVYT-RYM-L-TPE---D--R-----S---N-KQ- 698

FAP  692 DYLLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTN 741
CD26 699 E---------------Q----S----DVG--------T-ED--IASSTAH 748

FAP  742 *HLYTHMTHFLKQCFSLSD
CD26 749 Q-I----S--I------P
```

FIG. 2

	Breast Cancer	MFH	Healing Wound	Renal Cancer
FAPα	+ A	+ C	+ E	− G
CD26	− B	− D	+ F	+ H

Immunohistochemistry (See Kodachromes)

FIG. 3



ISOLATED PROTEINS CONTAINING PORTIONS OF FAPα AND OTHER PROTEINS

RELATED APPLICATION

This application is divisional of Ser. No. 08/940,391 filed Oct. 1, 1997 now U.S. Pat. No. 5,965,373 which a continuation-in-part of Ser. No. 08/230,491, filed Apr. 20, 1994 now U.S. Pat. No. 5,587,299.

FIELD OF THE INVENTION

This invention relates to certain molecules associated with cancer tissues and reactive tumor stromal cells. More particularly, it relates to fibroblast activation protein alpha ("FAPα" hereafter) molecules. A monomeric form of the molecule has previously been identified immunochemically, but nucleic acid molecules coding for it had not been isolated or cloned nor have dimers been identified. These, inter alia, are features of the invention. The monomeric protein has a molecular weight of from about 88 to about 95 kilodaltons as determined by SDS-PAGE of boiled samples. The dimer has a molecular weight of about 170 kilodaltons as determined by SDS-PAGE of unboiled samples. FAPα is characterized by a number of features and properties which are shared by and characteristic of membrane bound enzymes, suggesting very strongly that it, too, is a membrane bound enzyme. The nucleic acid molecules, which are a key part of the invention, are useful both as probes for cells expressing FAPα, and as starting materials for recombinant production of the protein. The FAPα protein can then be used to produce monoclonal antibodies specific for the protein and are thus useful diagnostic agents themselves. They also have additional uses, including uses related to enzymatic functions, as described herein.

BACKGROUND AND PRIOR ART

The invasive growth of epithelial cancers is associated with characteristic cellular and molecular changes in the supporting stroma. For example, epithelial cancers induce the formation of tumor blood vessels, the recruitment of reactive tumor stromal fibroblasts, lymphoid and phagocytic infiltrates, the release of peptide mediators and proteolytic enzymes, and the production of an altered extracellular matrix (ECM). See, e.g., Folkman, Adv. Cancer Res. 43: 175–203 (1985); Basset et al., Nature 348: 699–704 (1990); Denekamp et al., Cancer Metastasis Rev. 9: 267–282 (1990); Cullen et al., Cancer Res. 51: 4978–4985 (1991); Dvorak et al., Cancer Cells 3: 77–85 (1991); Liotta et al., Cancer Res. 51: 5054s-5059s (1991); Garin-Chesa et al., J. Histochem. Cytochem. 37: 1767–1776 (1989). A highly consistent molecular trait of the stroma in several common histologic types of epithelial cancers is induction of the fibroblast activation protein (FAPα), a cell surface glycoprotein with an observed M_r of 95,000 originally discovered with a monoclonal antibody, mAb F19, raised against proliferating cultured fibroblasts. See Rettig et al., Cancer Res. 46: 6406–6412 (1986); Rettig et al., Proc. Natl. Acad. Sci. USA 85: 3110–3114 (1988); Garin-Chesa et al., Proc. Natl. Acad. USA 87: 7235–7239 (1990); Rettig et al., Cancer Res. 53: 3327–3335 (1993). Each of these four papers is incorporated by reference in its entirety.

Immunohistochemical studies such as those cited supra have shown that FAPα is transiently expressed in certain normal fetal mesenchymal tissues but that normal adult tissues are generally FAPα$^-$. Similarly, malignant epithelial, neural and hematopoietic cells are generally FAPα$^-$. However, most of the common types of epithelial cancers, including >90% of breast, lung, skin, pancreas, and colorectal carcinomas, contain abundant FAPα$^+$ reactive stromal fibroblasts. Garin-Chesa et al., Proc. Natl. Acad. Sci. USA 87: 7235–7239 (1990). The FAPα$^+$ tumor stromal fibroblasts almost invariably accompany tumor blood vessels, forming a distinct cellular compartment interposed between the tumor capillary endothelium and the basal aspect of malignant epithelial cell clusters. While FAPα$^+$ stromal fibroblasts are found in both primary and metastatic carcinomas, benign and premalignant as epithelial lesions, such as fibroadenomas of the breast and colorectal adenomas only rarely contain FAPα$^+$ stromal cells. In contrast to the stroma-specific localization of FAPα in epithelial neoplasms, FAPα is expressed in the malignant cells of a large proportion of bone and soft tissue sarcomas. (Rettig et al., Proc. Natl. Acad. Sci. USA 85: 3110–3114 (1988)). Finally, FAPα$^+$ fibroblasts have been detected in the granulation tissue of healing wounds (Garin-Chesa et al., supra). Based on the restricted distribution pattern of FAPα in normal tissues and its uniform expression in the supporting stroma of many epithelial cancers, clinical trials with ^{131}I-labeled mAb F19 have been initiated in patients with metastatic colon cancer (Welt et al., Proc. Am. Assoc. Cancer Res. 33: 319 (1992); Welt et al. J. Clin. Oncol. 12: 1561–1571 (1994)) to explore the concept of "tumor stromal targeting" for immunodetection and immunotherapy of epithelial cancers.

Rettig et al., Int. J. Cancer 58: 385–392 (1994), incorporated by reference, discusses the FAPα molecule and its features. Rettig et al postulate that FAPα is found in high molecular weight complexes in excess of 400 kilodaltons, but do not discuss the possibility of dimeric molecules, nor does the paper elaborate on the specific enzymatic properties of the molecule.

The induction of FAPα$^+$ fibroblasts at times and sites of tissue remodeling during fetal development, tissue repair, and carcinogenesis is consistent with a fundamental role for this molecule in normal fibroblast physiology. Thus, it is of interest and value to isolate and to clone nucleic acid molecules which code for this molecule. This is one aspect of the invention, which is described in detail together with other features of the invention, in the disclosure which follows. Further aspects of the invention include the dimeric FAPα molecules, and the exploitation of the properties of these molecules. These features are also elaborated upon hereafter.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 compares the deduced amino acid sequence for FAPα, and the known sequence of CD26. The alignment has been optimized.

FIGS. 2A–2H, inclusive, display immunohistochemical detection of FAPα and CD26 in various tissues. In FIGS. 2A and 2B, breast cancer is studied, for FAPα (FIG. 2A), and CD26 (FIG. 2B). In FIGS. 2C and 2D, malignant fibrous histiocytoma is studied, for FAPα (FIG. 2C), and CD26 (FIG. 2D). Dermal scar tissue is examined in FIGS. 2E (FAPα), and 2F (CD26). Renal cell carcinoma is studied in FIGS. 2G (FAPα), and 2H (CD26).

FIG. 3 presents some of the data generated in experiments which showed that FAPα had extracellular matrix (ECM) protein degrading activity. When zymographic detection of gelatin degrading extracts of 293-FAP was carried out, the active substance was found to have a molecular weight of about 170 kD, via SDS-PAGE, using unboiled samples to preserve enzyme activity.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

EXAMPLE 1

Fibroblast cell line WI-38 had been observed, previously, to react with mAb F19 (Rettig et al., Canc. Res. 46: 6406–6412 (1986); Rettig et al., Proc. Natl. Acad. USA 85: 3110–3114 (1988); Garin-Chesa et al., Proc. Natl. Acad. Sci. USA 87: 7235–7239 (1990); Rettig et al., Canc. Res. 53: 3327–3335 (1993)). It was used in the experiments which follow.

A cDNA library was prepared from WI-38, using well known techniques and commercially available materials. Specifically, the library was constructed in expression vector pCDNAI, using the Fast Track mRNA isolation kit, and Librarian cDNA phagemid system. Once the library was prepared, the vectors were electroporated into cell line *E. coli* MC 1061/P3. The pCDNAI expression vector contains an antibiotic resistance gene, so the *E. coli* were selected via antibiotic resistance. The colonies which were resistant were then used in further experiments. The plasmid DNA from the colonies was obtained via alkaline lysis and purification on $CsCl_2$, in accordance with Sambrook et al, Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab, Cold Spring Harbor, N.Y. 2d Ed. 1989). The technique is well known to the art, but is incorporated by reference herein.

Once the plasmid DNA was isolated, it was used to transfect COS-1 cells, which were then cultured for forty-eight hours, after which these were tested with antibody coated dishes. The mAbs used included F19, as described by Rettig et al., (1986), supra, which is incorporated by reference in-its entirety. As COS-1 cells are normally FAPα$^-$, any positive results indicated the presence of the coding sequence. The immunoselection protocol was that of Aruffo et al., Proc. Natl. Acad. Sci USA 84: 3365–3369 (1987), incorporated by reference herein.

Plasmid DNA from positive clones was recovered, in accordance with Hirt, J. Mol. Biol. 26: 365–369 (1967), reintroduced into *E. coli* MC 1061/P3, and reselected in COS-1 cells.

The protocol presented herein was followed for four rounds. After this, the plasmid DNA of 50 isolated bacterial colonies was purified, using the Qiagen plasmid kit. Of the colonies, 27 clones were found to contain identical 2.8 kb inserts, as determined by EcoRI restriction enzyme mapping. Several of these were found to contain FAPα-specific cDNA as determined by transient expression in COS-1 cells and direct immunofluorescence staining with mAb F19. One of these clones, i.e., "pFAP.38" was selected for further study, as elaborated upon infra.

EXAMPLE 2

Once pFAP.38 had been identified, it was tested together with a vector coding for known cell surface marker CD26 ("pCD26"), as well as with control vector pCDNA I.

In these experiments, COS-1 cells were transfected with one of pFAP.38, pCD26, or pCDNAI. After forty-eight hours, the transfectants were tested, using the well known MHA rosetting assay for cell surface antigen expression. In these experiments, mAb F19, which is FAPα specific, was used, together with mAb EF-1, which is CD26 specific. Also used were four other FAPα specific mAbs, i.e., FB23, FB52, FB58 and C48. Also tested were two cancer cell lines, which are known to react with mAb F19 (SW872 liposarcoma), or EF-1 (SK-OV6 ovarian cancer). The results are set forth in Table 1, which follows.

TABLE 1

Cell surface expression of multiple FAPα epitopes and CD26 in human cells and COS-1 cell transfectants

	Cell surface antigen expression					
Target cell	F19	FB23	FB52	FB58	C48	EF-1
Human cells						
SW872 liposarcoma	>95%	>95%	>95%	>95%	>95%	—
SK-OV6 ovarian cancer	—	—	—	—	—	>95%
COS-1 transfectants						
COS . pCDNAI control	—	—	—	—	—	—
COS . pFAP 38	40%	30%	40%	20%	20%	—
COS . pCD26	—	—	—	—	—	40%

SEQ ID NO:10

EXAMPLE 3

Immunoprecipitation studies were then carried out to identify the antigen being targeted by the antibodies.

Cells were metabolically labelled with Trans ^{35}S-label, (ICN), extracted with lysis buffer (0.01M Tris-HCl/0.15M NaCl/0.01M $MgCl_2$/0.5% Nonidet P-40/aprotinin (20 ug/ml)/2 mM phenylmethyl-sulf onyl fluoride), and then immunoprecipitated. The protocols used are all well known, as will be seen by reference to Rettig et al., Canc. Res. 53: 3327–3335 (1993); and Fellinger et al., Canc. Res. 51: 336–340 (1991), the disclosures of which are all incorporated by reference in their entirety. Precipitating mAbs were negative control mouse Ig, mAb F19, or EF-1. Control tests were carried out with mock transfected COS-1 cells. Following immunoprecipitation, the immunoprecipitates were boiled in extraction buffer and separated by $NaDodSO_4$/PAGE, under reducing conditions. In some experiments, an additional test was carried out to determine whether or not the immunoprecipitated material was glycosylated. In these experiments, cell extracts were fractionated with Con A-SEPHAROSE prior to immunoprecipitation. Following immunoprecipitation, but prior to fractionation on $NaDodSO_4$/PAGE, these precipitates were digested with N-Glycanase.

The results showed that, in COS-1 cells, pFAP.38 directs expression of an 88 kd protein species (as determined via SDS-PAGE), which is slightly smaller than the 95 kd FAPα species produced by SW872, or cultured fibroblasts. Digestion with N-Glycanase produced peptides of comparable size (i.e., 74 kd versus 75 kd), showing that the glycosylation of the FAPA protein in COS-1 cells is different than in the human cell lines.

EXAMPLE 4

Classic Northern blot analysis was then carried out, using the mRNA from FAPα$^+$ fibroblast cell lines WI-38 and GM 05389, and FAPα$^-$ ovarian cancer cell line SK-OV6. Using the procedures of Sambrook et al., supra, five micrograms of mRNA from each cell line were tested. The probes used were ^{32}P labelled, and were prepared from a 2.3 kb ECO I fragment of pFAP.38, a 2.4 kb Hind III fragment of CD26, and a 1.8 kb BamHI fragment of γ-actin cDNA. These fragments had been purified from 1% agarose gels.

The extracts of FAPα$^+$ fibroblast strains showed a 2.8 kb FAP mRNA species, but extracts of SK-OV6 do not. A γ-actin mRNA species (1.8 kb), was observed in all species.

EXAMPLE 5

The cDNA identified as coding for FAPα was subjected to more detailed analysis, starting with sequencing. The classic Sanger methodology, as set forth in Proc. Natl. Acad. Sci. USA 74: 5463–5467 (1977), was used to sequence both strands of the cDNA. Once this was secured, an amino acid sequence was deduced therefrom. This information is presented in SEQ ID NO: 1. The sequence was then compared to the known amino acid sequence of CD26 (Morimoto et al., J. Immunol. 143: 3430–3437 (1989)). FIG. 1 presents the comparison, using optimized sequence alignment. Any gaps in the comparison are indicated by asterisks, while identical amino acids are shown by dashes in the CD26 sequence. A hydrophobic, putative transmembrane sequence is double underlined, while potential N-glycosylation sites are single underlined.

The sequence analysis shows a 2815 base pair insert, wherein 2280 base pairs constitute the open reading frame. This ORF extends from-start codon ATG at nucleotide 209, to stop codon TAA at 2486.

The deduced polypeptide is 760 amino acids long, and has a molecular weight of 87,832. In contrast, N-Glycanase digested, immunopurified FAPα was reported to have an estimated M_r of 75,000 on $NaDodSO_4$/PAGE (Rettig et al., Canc. Res. 53: 3327–3335 (1993)).

A GenBank data base search was carried out. The most closely related genes found were those encoding dipeptidyl peptidase IV homologues (DPPIV; EC 3.4.14.5), with human DPPIV (also known as T-cell activation antigen CD26), showing 51% nucleotide sequence identity, and 52% amino acid sequence identity.

The second set of related genes are human, rat, and bovine homologues of DPPX, a gene of unknown function widely expressed in brain and other normal tissues. The predicted human DPPX gene product shows about 30% amino acid sequence identity with FAPα and CD26. The FAPα molecule exhibits structural features typical of type II integral membrane proteins, including a large COOH-terminal extracellular domain, a hydrophobic transmembrane segment, and a short cytoplasmic tail. The putative extracellular domain contains five potential N-glycosylation sites, eleven cysteine residues (nine of which are conserved between FAPα and CD26), and three segments corresponding to highly conserved catalytic domains characteristic of serine proteases, such as DPPIV. These conserved sequences are presented in Table 2, which follows. Comparisons to DPPIV and DPPX were made via Morimoto et al., supra; Wada et al., Proc. Natl. Acad. Sci. USA 89: 197–201 (1992); Yokotani et al., Human Mol. Genet. 2: 1037–1039 (1993).

EXAMPLE 6

An additional set of experiments were carried out to determine whether FAPα related sequences are present in non-human species. To do so, human, mouse, and Chinese hamster genomic DNA was digested using restriction enzymes, and tested, via Southern blotting, using the 2.3 kb fragment, labelled with ^{32}P, described supra. Hybridization was carried out using stringent washing conditions (0.1× SSC, 0.1% $NaDodSO_4$, 68° C.). Cross-hybridization was readily observed with both the mouse and hamster DNA, suggesting the existence of highly conserved FAPα homologues. In control experiments using the CD26 cDNA fragment described supra, no evidence of cross hybridization was observed.

EXAMPLE 7

The CD26 molecule shares a number of biochemical and serological properties with FAPβ, which is a previously described, FAPα associated molecule having a molecular weight of 105 KD, and is found on cultured fibroblasts and melanocytes (Rettig et al., Canc. Res. 53: 3327–3335 (1993)). Cotransfection experiments were carried out to determine whether FAPβ is a CD26 gene product. To test this, the same protocols were used which were used for transfection with pFAP.38 or pCD26, as described supra, but using the two vectors. The results presented supra showed that cotransfection efficiency was about 40% for each vector, so about 10–20% of cell should be cotransfected.

Following cotransfection, the COS-1 cells were Trans ^{35}S-labeled, as described supra, then lysed, also as described supra.

The resulting cell extracts were separated on Con A SEPHAROSE, and the antigen (FAPα and/or CD26) were recovered in the Con A-bound fraction. The bound fraction was eluted with 0.25M α-D-mannopyranoside. Immunoprecipitation was then carried out, as described supra, and the precipitates were separated on $NaDodSO_4$/PAGE, also as discussed supra.

Those cells transfected only with pFAP.38 produced FAPα, but not FAPβ (determined from mAb F19 immunoprecipitates). They also produce no CD26 antigen (tested with EF-1). Those cells transfected with pCD26 alone produce CD26 but no FAPα. Cotransfectants produce CD26 and FAPα/FAPβ heteromers, as determined in the mAb F19 precipitates. This result provides direct evidence that FAPβ is a CD26 gene product.

EXAMPLE 8

It has been observed previously that some cultured human cell types coexpress FAPα and CD26, and show FAPα/

TABLE 2

Putative catalytic domains of FAPα, DPPIV and DPPX.

	624	702	734
Human FAPα	.... WGWSYGG SEQ ID NO:4	GTADDNV SEQ ID NO:6.	DQNHGLS SEQ ID NO:7 ....
Human DPPIV	.... WGWSYGG SEQ ID NO:4	GTADDNV SEQ ID NO:6	DEDHGIA SEQ ID NO:8 ....
Mouse DPPIV	.... WGWSYGG SEQ ID NO:4	GTADDNV SEQ ID NO:6	DEDHGIA SEQ ID NO:8 ....
Rat DPPIV	.... WGWSYGG SEQ ID NO:4	GTADDNV SEQ ID NO:6	DEDHGIA SEQ ID NO:8 ....
Yeast DPPIV	.... FGWSYGG SEQ ID NO:4	GTGDDNV SEQ ID NO:6	DSDHSIR SEQ ID NO:8 ....
Human DPPX	.... FGKDYGG SEQ ID NO:5	PTADEKI SEQ ID NO:9	DESHYFT SEQ ID NO:10 ....
Rat DPPX	.... FGKDYGG SEQ ID NO:5	ATADEKI SEQ ID NO:9	DESHYFH SEQ ID NO:10 ....
Bovine DPPX	.... FGKDYGG SEQ ID NO:5	ATADEKI SEQ ID NO:9	DESHYFS SEQ ID NO:10 ....

CD26 heteromer formation. In vivo distribution patterns of FAPα and CD26, however, as determined in previous immunohistochemical studies, appeared to be-non-overlapping. (See Rettig et al., Proc. Natl. Acad. Sci. USA85: 3110–3114 (1988); Garin-Chesa et al., Proc. Natl. Acad. Sci. USA 87: 7235–7329 (1990); Rettig et al., Canc. Res. 53: 3327–3335 (1993); Stein et al., in Knapp et al., eds. Leukocyte typing IV-white cell differentiation antigens, pp. 412–415 (Oxford University Press, N.Y. 1989), pp. 412–415; Möbious et al., J. Exp. Immunol. 74: 431–437 (1988)). In view of the potential significance of FAPα/CD26 coassociation, tissue distribution was reexamined, via side by side immunohistochemical staining of normal tissues and lesional tissues known to contain FAPAα$^+$ fibroblasts or FAPα$^+$ malignant cells.

To test the samples, they were embedded in OCT compound, frozen in isopentane precooled in liquid nitrogen, and stored at −70° C. until used. Five micrometer thick sections were cut, mounted on poly-L-lysine coated slides, air dried, and fixed in cold acetone (4° C., for 10 minutes). The sections were then tested with mAbs (10–20 ug/ml), using the well known avidin-biotin immmunoperoxidase method, as described by, e.g., Garin-Chesa et al., J. Histochem. Cytochem. 37: 1767–1776 (1989); Garin-Chesa et al., Proc. Natl. Acad. Sci. USA 87: 7235–7239 (1990); Rettig et al., Canc. Res. 53: 3327–3335 (1993); Garin-Chesa et al., Am. J. Pathol. 142: 557–567.

The results are shown in FIG. 2. Breast, colorectal, pancreas and lung carcinomas showed strong expression of FAPα and no CD26 was found (see FIGS. 2A and 2B). Five FAPα$^+$ sarcomas, including malignant fibrous histiocytoma (FIGS. 2C and 2D), were tested, and there was no expression of CD26. Examination of reactive fibroblasts of healing dermal wounds (FIGS. 2E, 2F), showed abundant expression of both FAPα and CD26. The three renal carcinomas tested (FIGS. 2G, 2H), showed expression of CD26 in malignant epithelium. FAPα was absent from malignant epithelial cells, and showed low expression in the stroma of these carcinomas.

EXAMPLE 9

A mammalian cell line, transfected with a FAPα encoding CDNA, was prepared.

Human embryonic kidney cell line 293 is well known and widely available from, e.g., the American Type Culture Collection.

Samples of 293 were maintained, in an incubator, at 37° C., in an atmosphere of 95% air, and 5% CO_2. The cells were cultured in a 50:50 mixture of Dulbecco's modified minimal essential medium and Ham's F12 medium, augmented with 10% fetal bovine serum, penicillin and streptomycin. Following the procedures described by Ustar et al., Eur. Mol. Biol. J. 1991, and/or Park et al., J. Biol. Chem. 169: 25646–25654 (1994), both of which are incorporated by reference, cDNA for FAPα (i.e., SEQ ID NO: 1), was transfected into the 293 cells. Details of the cDNA vector are provided, supra (pFAP.38). Transfectants were selected for resistance to antibiotics (200 ug/ml Geneticin), and were then maintained in selection medium, containing Geneticin.

Individual colonies of resistant cells were picked, grown to confluence in 6 well tissue culture plates, and were tested for FAPα expression in an immunofluorescence assay (IFA), using FAPα specific monoclonal antibody F19 as described supra.

Those colonies which expressed FAPα were expanded, and monitored by indirect IFA and cytofluorometric analysis, also as set forth, supra.

The IFAs were positive for the transfectants, referred to hereafter as cell line 293-FAP, but were negative for parental line 293.

EXAMPLE 10

In order to confirm that recombinant FAPα was, in fact, being produced, a series of immunoprecipitation experiments were carried out. These followed the methods of Park, et al., supra, and Rettig et al., Canc. Res. 53: 3327–3335 (1993), both of which are incorporated by reference. Essentially, ^{35}[S] methionine labelled cell extracts were combined with monoclonal antibody F19, in the manner described supra. Precipitates were then boiled in extraction buffer and run on SDS-PAGE gels, using, as a negative control, mouse IgGl. Both cell line 293-FAP, and non transfected line 293 were tested. The results indicated clearly, that recombinant FAPα was produced by the transfected cell line 293-FAP. This was determined by immunoprecipitation analyses, using FAPα specific monoclonal antibody F19.

EXAMPLE 11

The ability to produce recombinant FAPα permitted further study of the molecule's properties. Specifically, given the structural features outlined in the prior examples, experiments were designed to determine if FAPα possesses enzymatic activities. The experiments were designed to test whether or not FAPα had extracellular matrix (ECM) protein degrading activity.

Extracts of 293-FAP cells were prepared, using an extraction buffer (0.15M NaCl, 0.05M Tris-HCl, pH 7.4, 10 mM $MgCl_2$, 1 percent Triton X-114), were cleared by centrifugation (4,000×g, 10 minutes at 40° C.), and phase partitioned at 37° C. for 10–20 minutes. This was followed by further centrifugation (4000×g, 20 minutes at 20–25° C.) Detergent phases were diluted with buffer (0.15M NaCl, 0.05M Tris-HCl pH 7.4, 5 mM $CaCl_2$, 5 mM $MgCl_2$, 0.75% Empigen BB), and separated on concanavalin A-Sepharose following Rettig et al., supra. Any concanavalin A bound fractions were eluted with 0.25M methyl-α-D-mannopyranoside in elution buffer 0.15M NaCl,0.05M Tris-HCl, pH 7.4, 5 mM $CaCl_2$, 5 mM $MgCl_2$, 0.1% Triton X-100), mixed with zymography sample buffer (0.25M Tris-HCl, pH 6.8, 8% SDS, 40% glycerol, 0.01% bromophenol blue), at a 3:1 ratio, and used for further analysis.

Aliquots of sample were loaded onto polyacrylamide gels containing 0.1% of either of gelatin or casein. Electrophoresis was then carried out in a Biorad Mini-Protein II system, .at 20 mA constant current for 1.5–2 hours, until the bromophenol blue dye fronts of samples had reached the lower end of the gel. The gel was removed and incubated for one hour at 20–25° C. in a 2.5% aqueous solution of Triton X-100 on a rotary shaker. The Triton X-100 solution was decanted, and replaced with enzyme buffer (0.05M Tris-HCl, pH 7.5, 0.2M NaCl, 5 mM $CaCl_2$, 5 mM $MgCl_2$, 0.02% Brij 35). The gel was then incubated at 37° C. or 41° C., followed by staining or destaining at room temperature. Gels were stained with 0.5% of Coomassie Brilliant Blue G-250 in an aqueous solution of 30% methanol and 10% acetic acid for 15, 30, and 60 minutes, respectively. Subsequently, gels were incubated for 15 minutes in an aqueous solution of 30% CH_3OH and 5% glycerol, followed by drying between sheets of cellophane.

Gelatinase activity was evaluated in accordance with Kleiner et al., Anal. Biochem. 218: 325–329 (1994), incorporated by reference in its entirety. This is a routine assay

used to determine whether or not a protease capable of digesting gelatin is present. Labelled molecular weight standard were run on the same gels, under reducing conditions, for molecular weight determinations.

Proteolytic activity for defined amino acid sequence motifs were tested, using a well known membrane overlay assay. See Smith et al, Histochem. J. 24(9): 637–647 (1992), incorporated by reference. Substrates were Ala-Pro-7-amino-4-trifluoromethyl coumarin, Gly-Pro-7-amino-4-trifluorometbyl coumarin, and Lys-Pro-7-amino-4-trifluromethyl coumarin.

The results of these experiments are depicted, in part, in FIG. 3. This figure shows zymographic detection of gelatin degrading activity, in the cell extracts. See Kleiner et al., supra. A protein species of approximately 170 kilodaltons, as determined by SDS-PAGE, was observed to have gelatin degrading activity. This species, which was found in the 293-FAP cell line, but not in untransfected 293 cells, is thus identified as FAPα. The molecular weight is consistent with a dimer, i.e., a dimeric FAPα molecule.

The proteolytic activity described herein where gelatin is the substrate was not observed when casein was the substrate.

EXAMPLE 12

Further studies were then undertaken in order to characterize the 170 kD FAPα dimer further. Specifically, the experiments described in example 11 were repeated, except that 5% of 2-mercaptoethanol or 5 um iodoacetamide was added to the extracts prior to SDS-PAGE, or ethylenediamine N,N,N',N'-tetraacetic acid (10 mM) was added to the incubation buffer used for gelatin zymography. None of these treatments abolished the enzymatic activity. In contrast, heating at 100° C. for five minutes prior to SDS-polyacrylamide gel electrophoresis abolished the gelatin-degrading activity.

Further work, using a membrane overlay assay, described by, e.g., Smith et al., Histochem J. 24(9): 643–647 (1992), incorporated by reference, revealed that the FAPα dimers were able to cleave all of the Ala-Pro, Gly-Pro, and Lys-Pro dipeptides tested.

In further experiments, a fusion protein was produced which comprised the extracellular domains of both FAPα and murine CD8 proteins. This chimeric protein was produced in a baculovirus system in insect cells. The chimeric protein exhibited the same enzymatic activity as FAPα, using the model discussed supra.

The foregoing examples describe an isolated nucleic acid molecule which codes for fibroblast activating protein alpha ("FAPα"), as well as dimeric forms of the molecule, and uses thereof. The expression product of the sequence in COS-1 is a protein which, on SDS-PAGE of boiled samples, shows a molecular weight of about 88 kd. Deduced amino acid sequence, as provided in SEQ ID NO: 1, for one form of the molecule, yields a molecular weight of about 88 kd.

It should be noted that there is an apparent discrepancy in molecular weight in that the COS-1 isolate is glycosylated, while molecular weight from deduced amino acid sequences does not account for glycosylation. Membrane proteins are known to exhibit aberrant migration in gel systems, however, which may explain the difference observed here.

Also a part of the invention are chimeric and fusion proteins, which comprise a portion of FAPα which contain the molecule's catalytic domain, and additional, non FAPα components. The FAPα catalytic domain per se is also a part of the invention.

It is to be understood that, as described, FAPα may be glycosylated, with the type and amount of glycosylation varying, depending upon the type of cell expressing the molecule. The experiment described herein shows this. This is also true for the dimeric form of the molecule, first described herein, having a molecular weight of about 170 kilodaltons as determined by SDS-PAGE of unboiled samples.

The invention also comprehends the production of expression vectors useful in producing the FAPα molecule. In their broadest aspect, these vectors comprise the entire FAPα coding sequence or portions thereof, operably linked to a promoter. Additional elements may be a part of the expression vector, such as protein domains fused to the FAPα protein or protein portions ("fusion protein") genes which confer antibiotic resistance, amplifiable genes, and so forth.

The coding sequences and vectors may also be used to prepare cell lines, wherein the coding sequence or expression vector is used to transfect or to transform a recipient host. The type of cell used may be prokaryotic, such as *E. coli*, or eukaryotes, such as yeast, CHO, COS, or other cell types.

The identification of nucleic acid molecules such as that set forth in SEQ ID NO: 1 also enables the artisan to identify and to isolate those nucleic acid molecules which hybridize to it under stringent conditions. "Stringent condition" as used herein, refers to those parameters set forth supra, whereby both murine and hamster sequences were also identified. It will be recognized by the skilled artisan that these conditions afford a degree of stringency which can be achieved using parameters which vary from those recited. Such variance is apprehended by the expression "stringent conditions".

The ability of nucleic acid molecules to hybridize to complementary molecules also enables the artisan to identify cells which express FAPα, via the use of a nucleic acid hybridization assay. One may use the sequences described in the invention to hybridize to complementary sequences, and thus identify them. In this way, one can target mRNA, e.g., which is present in any cell expressing the FAPα molecule.

It is of course understood that the nucleic acid molecules of the invention are also useful in the production of recombinant FAPα, in both monomeric and dimeric form. The examples clearly show that host cells are capable of assembling the dimeric forms. The recombinant protein may be used, e.g., as a source of an immunogen for generation of antibodies akin to known mAb F19, and with the same uses. Similarly, the recombinant protein, and/or cells which express the molecule on their surface, may be used in assays to determine antagonists, agonists, or other molecules which interact with the FAPα molecule. Such molecules may be, but are not necessarily limited to, substrates, inhibiting molecules, antibodies, and so forth. This last feature of the invention should be considered in light of the observed structural resemblances to membrane bound enzymes. This type of molecule is associated with certain properties which need not be described in detail here. It will suffice to say that inhibition or potentiation of these properties as associated with FAPα is a feature of this invention. For example, one may identify substrates or the substrate for FAPα molecules, via the use of recombinant cells or recombinant FAPα per se. The substrates can be modified to improve their effect, to lessen their effect, or simply to label them with detectable signals so that they can be used, e.g., to identify cells which express FAPα. Study of the interaction of substrate and FAPα, as well as that between FAPα and any molecule

whatsoever, can be used to develop and/or to identify agonists and antagonists of the FAPα molecule.

Also a feature of the invention are isolated, dimeric FAPα molecules which have a molecular weight of about 170 kilodaltons as determined by SDS-PAGE, their use as an enzymatic cleaving agent, and other uses as are described herein. Enzymatically active forms of FAPα may also be produced as recombinant fusion proteins, comprising the catalytic domain of FAPα and other protein domains with suitable biochemical properties, including secretory signals protease cleavage sites, tags for purification, and other elements known to the artisan. The fact that FAPα has particular properties, as described herein, permits the identification of the molecule on cells expressing them. In turn, because the FAPα molecule is associated with tumors and tumor stromal cells, targeting of FAPα with therapeutic agents serves as a way to treat cancerous or precancerous condition, by administering sufficient therapeutic agent to alleviate cancer load.

The experiments showing the proteolytic properties of FAPα lead to yet a further aspect of the invention. It is well known that proteases which degrade extracellular matrix, or "ECM" proteins have an important role on certain aspects of tumor growth, including their effect on tumor cell invasion, tumor blood vessel formation (i.e., neoangiogenesis), and tumor metastasis. Collagens are of special interest vis-a-vis the substrates of proteases, as the collagens are an important part of the ECM. The fact that FAPα digests ECM suggests a therapeutic role for inhibitors of the molecule. "Inhibitors", as used herein, refers to molecules which interfere with FAPα enzyme function. Specifically excluded from such inhibitors is the monoclonal antibody F19. This mAb is known to bind to but not inhibit the enzyme function of FAPα, and hence it is not an inhibitor. The art is quite well versed with respect to monoclonal antibodies which both bind to and inhibit enzymes. Further examples of such inhibitors would include, e.g., substrate derivatives, such as modified collagen molecules, which interfere with the active site or sites of the FAPα molecule. Other suitable inhibitors will be apparent to the skilled artisan, and need not be listed here. In addition, the recombinant FAPα proteins and FAPα-transfected cell lines described supra can be employed in an enzymatic screening assay, using the substrate described supra or other suitable substrates, to identify inhibitors from any compound library. One can identify such enzyme inhibitors by combining a molecule which has FAP enzyme activity, such as the dimeric molecules of the invention, including dimers of SEQ ID NO: 2, with a substrate for the molecule with the enzymatic activity, as well as a substance believed to be an inhibitor. Then, one determines the activity of the molecule with enzymatic activity on its substrate, in the presence of the substance believed to be enzyme inhibitor. If there is a decrease in activity when the test substance is present as compared to when it is absent, then the substance is an inhibitor.

Other aspects of the invention will be clear to the skilled artisan, and need not be set forth here.

The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, it being recognized that various modifications are possible within the scope of the invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 10

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2815 Base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
AAGAACGCCC CCAAAATCTG TTTCTAATTT TACAGAAATC TTTTGAAACT TGGCACGGTA     60

TTCAAAAGTC CGTGGAAAGA AAAAAACCTT GTCCTGGCTT CAGCTTCCAA CTACAAAGAC    120

AGACTTGGTC CTTTTCAACG GTTTTCACAG ATCCAGTGAC CCACGCTCTG AAGACAGAAT    180

TAGCTAACTT TCAAAAAACAT CTGGAAAAAT GAAGACTTGG GTAAAAATCG TATTTGGAGT   240

TGCCACCTCT GCTGTGCTTG CCTTATTGGT GATGTGCATT GTCTTACGCC CTTCAAGAGT    300

TCATAACTCT GAAGAAAATA CAATGAGAGC ACTCACACTG AAGGATATTT TAAATGGAAC    360

ATTTTCTTAT AAAACATTTT TTCCAAACTG GATTTCAGGA CAAGAATATC TTCATCAATC    420

TGCAGATAAC AATATAGTAC TTTATAATAT TGAAACAGGA CAATCATATA CCATTTTGAG    480

TAATAGAACC ATGAAAAGTG TGAATGCTTC AAATTACGGC TTATCACCTG ATCGGCAATT    540

TGTATATCTA GAAAGTGATT ATTCAAAGCT TTGGAGATAC TCTTACACAG CAACATATTA    600
```

-continued

```
CATCTATGAC CTTAGCAATG GAGAATTTGT AAGAGGAAAT GAGCTTCCTC GTCCAATTCA    660
GTATTTATGC TGGTCGCCTG TTGGGAGTAA ATTAGCATAT GTCTATCAAA ACAATATCTA    720
TTTGAAACAA AGACCAGGAG ATCCACCTTT TCAAATAACA TTTAATGGAA GAGAAAATAA    780
AATATTTAAT GGAATCCCAG ACTGGGTTTA TGAAGAGGAA ATGCTTCCTA CAAAATATGC    840
TCTCTGGTGG TCTCCTAATG GAAAATTTTT GGCATATGCG GAATTTAATG ATAAGGATAT    900
ACCAGTTATT GCCTATTCCT ATTATGGCGA TGAACAATAT CCTAGAACAA TAAATATTCC    960
ATACCCAAAG GCTGGAGCTA AGAATCCCGT TGTTCGGATA TTTATTATCG ATACCACTTA   1020
CCCTGCGTAT GTAGGTCCCC AGGAAGTGCC TGTTCCAGCA ATGATAGCCT CAAGTGATTA   1080
TTATTTCAGT TGGCTCACGT GGGTTACTGA TGAACGAGTA TGTTTGCAGT GGCTAAAAAG   1140
AGTCCAGAAT GTTTCGGTCC TGTCTATATG TGACTTCAGG GAAGACTGGC AGACATGGGA   1200
TTGTCCAAAG ACCCAGGAGC ATATAGAAGA AAGCAGAACT GGATGGGCTG GTGGATTCTT   1260
TGTTTCAAGA CCAGTTTTCA GCTATGATGC CATTTCGTAC TACAAAATAT TTAGTGACAA   1320
GGATGGCTAC AAACATATTC ACTATATCAA AGACACTGTG GAAAATGCTA TTCAAATTAC   1380
AAGTGGCAAG TGGGAGGCCA TAAATATATT CAGAGTAACA CAGGATTCAC TGTTTTATTC   1440
TAGCAATGAA TTTGAAGAAT ACCCTGGAAG AAGAAACATC TACAGAATTA GCATTGGAAG   1500
CTATCCTCCA AGCAAGAAGT GTGTTACTTG CCATCTAAGG AAAGAAAGGT GCCAATATTA   1560
CACAGCAAGT TTCAGCGACT ACGCCAAGTA CTATGCACTT GTCTGCTACG GCCCAGGCAT   1620
CCCCATTTCC ACCCTTCATG ATGGACGCAC TGATCAAGAA ATTAAAATCC TGGAAGAAAA   1680
CAAGGAATTG GAAAATGCTT TGAAAAATAT CCAGCTGCCT AAAGAGGAAA TTAAGAAACT   1740
TGAAGTAGAT GAAATTACTT TATGGTACAA GATGATTCTT CCTCCTCAAT TTGACAGATC   1800
AAAGAAGTAT CCCTTGCTAA TTCAAGTGTA TGGTGGTCCC TGCAGTCAGA GTGTAAGGTC   1860
TGTATTTGCT GTTAATTGGA TATCTTATCT TGCAAGTAAG GAAGGGATGG TCATTGCCTT   1920
GGTGGATGGT CGAGGAACAG CTTTCCAAGG TGACAAACTC CTCTATGCAG TGTATCGAAA   1980
GCTGGGTGTT TATGAAGTTG AAGACCAGAT TACAGCTGTC AGAAAATTCA TAGAAATGGG   2040
TTTCATTGAT GAAAAAAGAA TAGCCATATG GGGCTGGTCC TATGGAGGAT ACGTTTCATC   2100
ACTGGCCCTT GCATCTGGAA CTGGTCTTTT CAAATGTGGT ATAGCAGTGG CTCCAGTCTC   2160
CAGCTGGGAA TATTACGCGT CTGTCTACAC AGAGAGATTC ATGGGTCTCC CAACAAAGGA   2220
TGATAATCTT GAGCACTATA AGAATTCAAC TGTGATGGCA AGAGCAGAAT ATTTCAGAAA   2280
TGTAGACTAT CTTCTCATCC ACGGAACAGC AGATGATAAT GTGCACTTTC AAAACTCAGC   2340
ACAGATTGCT AAAGCTCTGG TTAATGCACA AGTGGATTTC CAGGCAATGT GGTACTCTGA   2400
CCAGAACCAC GGCTTATCCG GCCTGTCCAC GAACCACTTA TACACCCACA TGACCCACTT   2460
CCTAAAGCAG TGTTTCTCTT TGTCAGACTA AAAACGATGC AGATGCAAGC CTGTATCAGA   2520
ATCTGAAAAC CTTATATAAA CCCCTCAGAC AGTTTGCTTA TTTTATTTTT TATGTTGTAA   2580
AATGCTAGTA TAAACAAACA AATTAATGTT GTTCTAAAGG CTGTTAAAAA AAAGATGAGG   2640
ACTCAGAAGT TCAAGCTAAA TATTGTTTAC ATTTTCTGGT ACTCTGTGAA AGAAGAGAAA   2700
AGGGAGTCAT GCATTTTGCT TTGGACACAG TGTTTTATCA CCTGTTCATT TGAAGAAAAA   2760
TAATAAAGTC AGAAGTTCAA AAAAAAAAAA AAAAAAAAAA AAAGCGGCCG CTCGA        2815
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 760 amino acids

-continued

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Lys Thr Trp Val Lys Ile Val Phe Gly Val Ala Thr Ser Ala Val
5                   10                  15

Leu Ala Leu Leu Val Met Cys Ile Val Leu Arg Pro Ser Arg Val His
20                  25                  30

Asn Ser Glu Glu Asn Thr Met Arg Ala Leu Thr Leu Lys Asp Ile Leu
35                  40                  45

Asn Gly Thr Phe Ser Tyr Lys Thr Phe Phe Pro Asn Trp Ile Ser Gly
50                  55                  60

Gln Glu Tyr Leu His Gln Ser Ala Asp Asn Asn Ile Val Leu Tyr Asn
65                  70                  75                  80

Ile Glu Thr Gly Gln Ser Tyr Thr Ile Leu Ser Asn Arg Thr Met Lys
85                  90                  95

Ser Val Asn Ala Ser Asn Tyr Gly Leu Ser Pro Asp Arg Gln Phe Val
100                 105                 110

Tyr Leu Glu Ser Asp Tyr Ser Lys Leu Trp Arg Tyr Ser Tyr Thr Ala
115                 120                 125

Thr Tyr Tyr Ile Tyr Asp Leu Ser Asn Gly Glu Phe Val Arg Gly Asn
130                 135                 140

Glu Leu Pro Arg Pro Ile Gln Tyr Leu Cys Trp Ser Pro Val Gly Ser
145                 150                 155                 160

Lys Leu Ala Tyr Val Tyr Gln Asn Asn Ile Tyr Leu Lys Gln Arg Pro
165                 170                 175

Gly Asp Pro Pro Phe Gln Ile Thr Phe Asn Gly Arg Glu Asn Lys Ile
180                 185                 190

Phe Asn Gly Ile Pro Asp Trp Val Tyr Glu Glu Glu Met Leu Pro Thr
195                 200                 205

Lys Tyr Ala Leu Trp Trp Ser Pro Asn Gly Lys Phe Leu Ala Tyr Ala
210                 215                 220

Glu Phe Asn Asp Lys Asp Ile Pro Val Ile Ala Tyr Ser Tyr Tyr Gly
225                 230                 235                 240

Asp Glu Gln Tyr Pro Arg Thr Ile Asn Ile Pro Tyr Pro Lys Ala Gly
245                 250                 255

Ala Lys Asn Pro Val Val Arg Ile Phe Ile Ile Asp Thr Thr Tyr Pro
260                 265                 270

Ala Tyr Val Gly Pro Gln Glu Val Pro Val Pro Ala Met Ile Ala Ser
275                 280                 285

Ser Asp Tyr Tyr Phe Ser Trp Leu Thr Trp Val Thr Asp Glu Arg Val
290                 295                 300

Cys Leu Gln Trp Leu Lys Arg Val Gln Asn Val Ser Val Leu Ser Ile
305                 310                 315                 320

Cys Asp Phe Arg Glu Asp Trp Gln Thr Trp Asp Cys Pro Lys Thr Gln
325                 330                 335

Glu His Ile Glu Glu Ser Arg Thr Gly Trp Ala Gly Gly Phe Phe Val
340                 345                 350

Ser Arg Pro Val Phe Ser Tyr Asp Ala Ile Ser Tyr Tyr Lys Ile Phe
355                 360                 365

Ser Asp Lys Asp Gly Tyr Lys His Ile His Tyr Ile Lys Asp Thr Val
370                 375                 380

Glu Asn Ala Ile Gln Ile Thr Ser Gly Lys Trp Glu Ala Ile Asn Ile
385                 390                 395                 400
```

Phe Arg Val Thr Gln Asp Ser Leu Phe Tyr Ser Ser Asn Glu Phe Glu
405 410 415

Glu Tyr Pro Gly Arg Arg Asn Ile Tyr Arg Ile Ser Ile Gly Ser Tyr
420 425 430

Pro Pro Ser Lys Lys Cys Val Thr Cys His Leu Arg Lys Glu Arg Cys
435 440 445

Gln Tyr Tyr Thr Ala Ser Phe Ser Asp Tyr Ala Lys Tyr Tyr Ala Leu
450 455 460

Val Cys Tyr Gly Pro Gly Ile Pro Ile Ser Thr Leu His Asp Gly Arg
465 470 475 480

Thr Asp Gln Glu Ile Lys Ile Leu Glu Glu Asn Lys Glu Leu Glu Asn
485 490 495

Ala Leu Lys Asn Ile Gln Leu Pro Lys Glu Glu Ile Lys Lys Leu Glu
500 505 510

Val Asp Glu Ile Thr Leu Trp Tyr Lys Met Ile Leu Pro Pro Gln Phe
515 520 525

Asp Arg Ser Lys Lys Tyr Pro Leu Leu Ile Gln Val Tyr Gly Gly Pro
530 535 540

Cys Ser Gln Ser Val Arg Ser Val Phe Ala Val Asn Trp Ile Ser Tyr
545 550 555 560

Leu Ala Ser Lys Glu Gly Met Val Ile Ala Leu Val Asp Gly Arg Gly
565 570 575

Thr Ala Phe Gln Gly Asp Lys Leu Leu Tyr Ala Val Tyr Arg Lys Leu
580 585 590

Gly Val Tyr Glu Val Glu Asp Gln Ile Thr Ala Val Arg Lys Phe Ile
595 600 605

Glu Met Gly Phe Ile Asp Glu Lys Arg Ile Ala Ile Trp Gly Trp Ser
610 615 620

Tyr Gly Gly Tyr Val Ser Ser Leu Ala Leu Ala Ser Gly Thr Gly Leu
625 630 635 640

Phe Lys Cys Gly Ile Ala Val Ala Pro Val Ser Ser Trp Glu Tyr Tyr
645 650 655

Ala Ser Val Tyr Thr Glu Arg Phe Met Gly Leu Pro Thr Lys Asp Asp
660 665 670

Asn Leu Glu His Tyr Lys Asn Ser Thr Val Met Ala Arg Ala Glu Tyr
675 680 685

Phe Arg Asn Val Asp Tyr Leu Leu Ile His Gly Thr Ala Asp Asp Asn
690 695 700

Val His Phe Gln Asn Ser Ala Gln Ile Ala Lys Ala Leu Val Asn Ala
705 710 715 720

Gln Val Asp Phe Gln Ala Met Trp Tyr Ser Asp Gln Asn His Gly Leu
725 730 735

Ser Gly Leu Ser Thr Asn His Leu Tyr Thr His Met Thr His Phe Leu
740 745 750

Lys Gln Cys Phe Ser Leu Ser Asp
755 760

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 766 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

-continued

```
Met Lys Thr Pro Trp Lys Val Leu Leu Gly Leu Leu Gly Ala Ala Ala
                5                   10                  15

Leu Val Thr Ile Ile Thr Val Pro Val Val Leu Leu Asn Lys Gly Thr
            20                  25                  30

Asp Asp Ala Thr Ala Asp Ser Arg Lys Thr Tyr Thr Leu Thr Asp Tyr
        35                  40                  45

Leu Lys Asn Thr Tyr Arg Leu Lys Leu Tyr Ser Leu Arg Trp Ile Ser
    50                  55                  60

Asp His Glu Tyr Leu Tyr Lys Gln Glu Asn Asn Ile Leu Val Phe Asn
65                  70                  75                  80

Ala Glu Tyr Gly Asn Ser Ser Val Phe Leu Glu Asn Ser Thr Phe Asp
                85                  90                  95

Glu Phe Gly His Ser Ile Asn Asp Tyr Ser Ile Ser Pro Asp Gly Gln
            100                 105                 110

Phe Ile Leu Leu Glu Tyr Asn Tyr Val Lys Gln Trp Arg His Ser Tyr
        115                 120                 125

Thr Ala Ser Tyr Asp Ile Tyr Asp Leu Asn Lys Arg Gln Leu Ile Thr
    130                 135                 140

Glu Glu Arg Ile Pro Asn Asn Thr Gln Trp Val Thr Trp Ser Pro Val
145                 150                 155                 165

Gly His Lys Leu Ala Tyr Val Trp Asn Asn Asp Ile Tyr Val Lys Ile
                170                 175                 180

Glu Pro Asn Leu Pro Ser Tyr Arg Ile Thr Trp Thr Gly Lys Glu Asp
            185                 190                 195

Ile Ile Tyr Asn Gly Ile Thr Asp Trp Val Tyr Glu Glu Glu Val Phe
        200                 205                 210

Ser Ala Tyr Ser Ala Leu Trp Trp Ser Pro Asn Gly Thr Phe Leu Ala
    215                 220                 225

Tyr Ala Gln Phe Asn Asp Thr Glu Val Pro Leu Ile Glu Tyr Ser Phe
230                 235                 240                 245

Tyr Ser Asp Glu Ser Leu Gln Tyr Pro Lys Thr Val Arg Val Pro Tyr
                250                 255                 260

Pro Lys Ala Gly Ala Val Asn Pro Thr Val Lys Phe Phe Val Val Asn
            265                 270                 275

Thr Asp Ser Leu Ser Ser Val Thr Asn Ala Thr Ser Ile Gln Ile Thr
        280                 285                 290

Ala Pro Ala Ser Met Leu Ile Gly Asp His Tyr Leu Cys Asp Val Thr
    295                 300                 305

Trp Ala Thr Gln Glu Arg Ile Ser Leu Gln Trp Leu Arg Arg Ile Gln
310                 315                 320                 325

Asn Tyr Ser Val Met Asp Ile Cys Asp Tyr Asp Glu Ser Ser Gly Arg
                330                 335                 340

Trp Asn Cys Leu Val Ala Arg Gln His Ile Glu Met Ser Thr Thr Gly
            345                 350                 355

Trp Val Gly Arg Phe Arg Pro Ser Glu Pro His Phe Thr Leu Asp Gly
        360                 365                 370

Asn Ser Phe Tyr Lys Ile Ile Ser Asn Glu Glu Gly Tyr Arg His Ile
    375                 380                 385

Cys Tyr Phe Gln Ile Asp Lys Lys Asp Cys Thr Phe Ile Thr Lys Gly
390                 395                 400                 405

Thr Trp Glu Val Ile Gly Ile Glu Ala Leu Thr Ser Asp Tyr Leu Tyr
                410                 415                 420

Tyr Ile Ser Asn Glu Tyr Lys Gly Met Pro Gly Gly Arg Asn Leu Tyr
```

-continued

```
            425                 430                 435

Lys Ile Gln Leu Ser Asp Tyr Thr Lys Val Thr Cys Leu Ser Cys Glu
        440                 445                 450

Leu Asn Pro Glu Arg Cys Gln Tyr Tyr Ser Val Ser Phe Ser Lys Glu
    455                 460                 460

Ala Lys Tyr Tyr Gln Leu Arg Cys Ser Gly Pro Gly Leu Pro Leu Tyr
465                 470                 475                 480

Thr Leu His Ser Ser Val Asn Asp Lys Gly Leu Arg Val Leu Glu Asp
                485                 490                 495

Asn Ser Ala Leu Asp Lys Met Leu Gln Asn Val Gln Met Pro Ser Lys
            500                 505                 510

Lys Leu Asp Phe Ile Ile Leu Asn Glu Thr Lys Phe Trp Tyr Gln Met
        515                 520                 525

Ile Leu Pro Pro His Phe Asp Lys Ser Lys Lys Tyr Pro Leu Leu Leu
    530                 535                 540

Asp Val Tyr Ala Gly Pro Cys Ser Gln Lys Ala Asp Thr Val Phe Arg
545                 550                 555                 560

Leu Asn Trp Ala Thr Tyr Leu Ala Ser Thr Glu Asn Ile Ile Val Ala
                565                 570                 575

Ser Phe Asp Gly Arg Gly Ser Gly Tyr Gln Gly Asp Lys Ile Met His
            580                 585                 590

Ala Ile Asn Arg Arg Leu Gly Thr Phe Glu Val Glu Asp Gln Ile Glu
        595                 600                 605

Ala Ala Arg Gln Phe Ser Lys Met Gly Phe Val Asp Asn Lys Arg Ile
    610                 615                 620

Ala Ile Trp Gly Trp Ser Tyr Gly Gly Tyr Val Thr Ser Met Val Leu
625                 630                 635                 640

Gly Ser Gly Ser Gly Val Phe Lys Cys Gly Ile Ala Val Ala Pro Val
                645                 650                 655

Ser Arg Trp Glu Tyr Tyr Asp Ser Val Tyr Thr Glu Arg Tyr Met Gly
            660                 665                 670

Leu Pro Thr Pro Glu Asp Asn Leu Asp His Tyr Arg Asn Ser Thr Val
        675                 680                 685

Met Ser Arg Ala Glu Asn Phe Lys Gln Val Glu Tyr Leu Leu Ile His
    690                 695                 700

Gly Thr Ala Asp Asp Asn Val His Phe Gln Gln Ser Ala Gln Ile Ser
705                 710                 715                 720

Lys Ala Leu Val Asp Val Gly Val Asp Phe Gln Ala Met Trp Tyr Thr
                725                 730                 735

Asp Glu Asp His Gly Ile Ala Ser Ser Thr Ala His Gln His Ile Tyr
            740                 745                 750

Thr His Met Ser His Phe Ile Lys Gln Cys Phe Ser Leu Pro
        755                 760                 765
```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: The first Xaa is either Trp or Phe.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Xaa Gly Trp Ser Tyr Gly Gly

5

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Phe Gly Lys Asp Tyr Gly Gly
5

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: Xaa is either Ala or Gly

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Gly Thr Xaa Asp Asp Asn Val
5

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Ala Gln Asn His Gly Leu Ser
5

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION:
The first Xaa is Glu or Ser. When the first Xaa is Glu, the second Xaa is Gly and the third is Ala. When the first Xaa is Ser, the second Xaa is Ser, and the third Xaa is Arg.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Asp Xaa Asp His Xaa Ile Xaa
5

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION:
Xaa is Pro or Ala.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Xaa Thr Ala Asp Glu Lys Ile
5

-continued

```
(2) INFORMATION FOR SEQ ID NO: 10:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH:   7 amino acids
          (B) TYPE:     amino acid
          (D) TOPOLOGY: linear

    (ix) FEATURE:
          (D) OTHER INFORMATION:
               Xaa is Thr, His or Ser.

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Asp Glu Ser His Tyr Phe Xaa
                5
```

We claim:

1. An isolated protein consisting of:

(i) an extracellular domain of an FAPα protein, the amino acid sequence of which is set forth in SEQ ID NO: 1, and

(ii) at least one domain of a non FAPα protein.

2. The Isolated protein of claim 1, wherein said non FAPα protein is a CD8 protein.

3. The isolated protein of claim 1, wherein said at least one domain of a non FAPα protein is an extracellular domain of a CD8 protein.

4. The isolated protein of claim 1, wherein said protein is a chimeric protein.

5. The isolated protein of claim 1, wherein said protein is a fusion protein.

* * * * *